DOCUMENT RESUME

ED 385 553 TM 023 970

AUTHOR Murphy, Richard T.
TITLE Educational Effectiveness of Sesame Street: A Review of the First Twenty Years of Research, 1969-1989.
INSTITUTION Educational Testing Service, Princeton, N.J.
SPONS AGENCY Children's Television Workshop, New York, N.Y.
REPORT NO ETS-RR-91-55
PUB DATE Oct 91
NOTE 86p.
PUB TYPE Information Analyses (070)

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS *Childrens Television; Educational History; Educational Objectives; Educational Research; Educational Television; Foreign Countries; *Instructional Effectiveness; Literature Reviews; *Mass Media Effects; *Preschool Education; Reading Readiness; Research Design; Research Methodology; *School Readiness; Synthesis
IDENTIFIERS Australia; Canada; Israel; Mexico; *Sesame Street

ABSTRACT

Research studies that have focused on the educational effectiveness of "Sesame Street" are reviewed, and a summary and synthesis of research results are presented. Educational effectiveness is defined as effectiveness in areas related to beginning schooling and the early years of the child's formal education. From the more than 100 empirical studies of the impact of "Sesame Street," 16 were chosen because they collected data on individual children. On the whole, the studies with the strongest designs indicate that "Sesame Street" had a significant positive impact on the children in terms of the variables measured and relative to the children in other groups studied. In addition, the effect was large enough to have shown up in studies with very small samples. The research studies reviewed in this report indicate that "Sesame Street" has had a significant positive impact on the pre-reading and school-readiness skills of children in the United States, Australia, Canada, Israel, and Mexico. An appendix contains summaries of all 16 studies. (Contains 35 references.) (SLD)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

RR-91-55

ED 385 553

RESEARCH REPORT

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

H. I. BRAUN

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

# EDUCATIONAL EFFECTIVENESS OF SESAME STREET: A REVIEW OF THE FIRST TWENTY YEARS OF RESEARCH 1969 – 1989

Richard T. Murphy

Educational Testing Service
Princeton, New Jersey
October 1991

TM023970

BEST COPY AVAILABLE

# Educational Effectiveness of Sesame Street:

## A Review of the First Twenty Years of Research*

## 1969 - 1989

Richard T. Murphy

* Children's Television Workshop funded the research upon which this report is based.

Copyright © 1991. Educational Testing Service. All rights reserved.

## Table of Contents

**Acknowledgements**

**Abstract**

**Section 1: Study Selection and Methodology** ..... 1

**Section 2: Characteristics and Results of the Studies** ..... 6

**Section 3: Synthesis of Research Results** ..... 15

**References** ..... 28

**Appendix: Summaries of Research Studies**

## Acknowledgements

Children's Television Workshop (CTW) supported this review as part of their commemoration of the 20th anniversary of **Sesame Street**. They also provided access to their offices in New York City and to copies of many of the basic documents available on **Sesame Street** studies conducted during the last twenty years. Richard Luker (University of Pennsylvania), Jan Fernback (University of Pennsylvania), and Gloria Sammur (CTW) compiled a comprehensive research bibliography and provided abstracts for most of the entries. Their work facilitated this review.

I am also grateful to Gita Wilder and Lori Morris, colleagues at Educational Testing Service (ETS), for comments on a preliminary draft of this report. They are responsible for making the report more readable and more accurate. Finally, thanks to Adele Lechowicz, Senior Administrative Assistant at ETS, for preparing the manuscript and tables.

The interpretations of research results and the conclusions drawn in the review are mine. I welcome comments on the report and any suggestions for conducting further reviews of research.

Richard T. Murphy
Senior Research Psychologist
Educational Testing Service

## Abstract

In this report, we review **Sesame Street** research studies that focused on the educational effectiveness of the series and provide a summary and synthesis of the results of the research. By educational effectiveness, we mean effectiveness in those areas related to beginning schooling, to the early years of children's formal education. An important goal of **Sesame Street** was to help children prepare for their initial school experience by exposing them to a variety of rich and appealing examples. Studies that attempt to provide evidence about the successes and/or failures of **Sesame Street** in that effort are the focus of this review.

## Section 1: Study Selection and Methodology

The purpose of this research review is to summarize and synthesize the results obtained and reported by researchers who have studied the effects of **Sesame Street** on its young viewers over the past twenty years. The review was conducted as part of the commemoration of the 20th anniversary of **Sesame Street**'s first broadcast in November 1969. A comprehensive bibliography of the literature on **Sesame Street** was prepared by Dr. Richard Luker, University of Pennsylvania, and a group of researchers working under his direction. Abstracts of the entries were written and copies of the publications and research documents were gathered together. In conducting this review, ETS staff made extensive use of the research bibliography, abstracts, and documents gathered by Dr. Luker and his associates.

The **Sesame Street Research Bibliography** (1989) contains more than one hundred citations for empirical studies of the educational impact of **Sesame Street**. The actual publications and the basic documents on which the publications are based are available at the Children's Television Workshop offices in New York City and were made available to ETS staff. In reviewing those materials, we found that a number of the citations were based on the same underlying studies. There were not as many independent studies of educational impact as we thought there might be at first. In the ETS review, we included all studies that collected data on individual children. The only criteria for inclusion in the review was that the authors provided data on some educationally related variable. In all, we identified the 16 independent studies which are reviewed in the following sections of this report.

In the process of selecting a methodology for this review, ETS staff examined current reviews of research and identified three general approaches to summarizing research results: (1) a "Meta-analysis" approach (Glass, 1976; Glass and Smith, 1976), (2) a "Summing Up" approach (Light and Pillemer, 1984), and (3) a "Best Evidence Synthesis" approach (Slavin, 1986).

The "Meta-analysis" approach requires the computation of a standardized "effect size," i.e., a measure of the value added to or subtracted from a treatment group's performance due to exposure to the treatment, for each study. The "effect sizes" are then analyzed and aggregated to give an overall effect for the program or treatment being studied. Because the approach attempts to even out differences across studies, the method requires a relatively large number of studies for its appropriate use. In examining reviews that used a "Meta-analysis" approach in two areas, i.e., computer-assisted instruction (Kulik, Kulik, & Cohen, 1980) and elementary school reading programs (Slavin & Madden, 1989), we noted a tendency of reviewers to give only very brief attention to complex studies with multiple sites, or to eliminate them completely from the review. For that reason, and also because the number of studies in this review was small, it was decided that "Meta-analysis" was not an appropriate methodology for this review. However, effect sizes for the studies examined were calculated and are given in the tables in the following sections.

"Summing Up" is the phrase used to describe a comprehensive approach to research synthesis described by Light and Pillemer (1984). They investigated the issue of synthesizing research results and suggested in their **Summing Up: The Science of Reviewing Research** a combination of narrative and quantitative techniques built around a number of important issues such as the **Precise Purpose of the Review**, the **Studies Selected**, the **Similarity of Treatments and Control/Comparison Groups**, **Publication Bias**, **Outcomes and Research Design**, and **Units of Analyses across Studies**. Although we have not used the Light and Pillemer approach explicitly, we agree with the issues they raise and treat them in the narrative reviews of the studies.

Finally, the "Best Evidence Synthesis" approach is still another approach which encourages the use of both quantitative and qualitative techniques to make the case for an interpretable synthesis. Slavin (1986) surveyed a large number of research syntheses and noted that there were cases in which important, well designed studies that used complex designs, ANCOVAs, and multiple regression

analyses did not provide sufficient information to allow for computation of effect sizes and were excluded from the synthesis. He recommended an alternative method to that of "Meta-analysis," an approach which he called a "Best Evidence Synthesis." In that approach, Slavin suggested that some studies be integrated using a "Meta-analysis" approach and others be treated in a more narrative fashion. The suggestion of Slavin is a useful one, though, as Slavin points out, there remains the problem of defining "best evidence." In this review, an approach similar to that suggested by Slavin is used. The rationale for "best-evidence" is based on a framework involving the estimation of causal effects suggested by Holland and Rubin (Rubin, 1974; Holland and Rubin, 1983; Holland, 1986; Holland and Rubin, 1987) and an approach involving degrees of control as implemented in the research review conducted by Messick (1980) in a review of the research on the effects of coaching on student performance on the Scholastic Aptitude Test (SAT).

In this **Sesame Street** review, "best" is judged in terms of the degree to which the results and the evidence fulfill criteria set by Holland and Rubin, i.e., by the extent to which the study results provide evidence that a significant effect has occurred and that the effect can be attributed to the treatment or program being studied. Holland and Rubin, in the references cited above, stress the importance of clarifying and making explicit the relationships that exist between the components of a research study and the subsequent conclusions that an "effect" was "caused" by the treatment in the research study. They define an **effect** as the difference between the **behavior of the experimental group in the treatment condition** and the generally unobservable **behavior of the same treatment group in the control condition**. By being precise and explicit about the assumptions underlying the strategy used for estimating how the treatment group would have performed in the absence of the treatment, researchers, Holland and Rubin suggest, can improve the validity of their arguments and conclusions. Basically, the Holland and Rubin suggest that the strength of the evidence is based on the credibility of the estimate that is used for the performance of the treatment group in the control condition. If strong assumptions are needed to

justify the estimate, then the credibility of the argument is weakened. The best evidence is the evidence that reinforces the assertion that the effect is due to the treatment and that other plausible causes of the effect can be ruled out.

The following brief technical explanation of the approach is adapted from Holland and Rubin (1983) and given for those who are interested. The basic elements of the model are given below:

- **A population of units, P;**
- **An "experimental manipulation" with t and c conditions;**
- **An outcome variable, Y; and**
- **A concomitant variable, X.**

Holland and Rubin define the *causal effect* of t on Y (relative to c) for each unit in P as the difference $Y_t - Y_c$. This is the amount that t (the treatment condition) has increased (or decreased) the value of Y (relative to c) on each unit. The expected value $E(Y_t - Y_c)$ is the average causal effect of t versus c on Y in P. The average difference is equivalent to the average value of Y for units in T minus the average value of Y for units in T when exposed to C. But, when dealing with students, units cannot be exposed to both T and C. This is the central problem. Any analysis can only estimate $Y_c$ for units in T. What would happen to the experimental treatment children in the control condition can never be known directly. The process used and the assumptions made by a researcher in estimating $Y_c$ and $Y_t - Y_c$ is the critical component that differentiates one set of results from another in the area of analysis and the use of inference to derive conclusions. In the Holland and Rubin model, the concomitant variable X is used to provide a practical, credible estimate of $Y_c$.

Holland and Rubin demonstrate how researchers can reach contradictory conclusions from the same data if they base their analyses on different assumptions. In the review, we demonstrate instances in which conclusions drawn by researchers are actually based on unstated assumptions with little credibility. And, in other instances, we demonstrate how arguments are strengthened by the presentation of evidence which supports initially weak assumptions. Finally,

different unstated assumptions are not the only dangers to interpretation of research results. There may also be differences in the populations included in the studies, differences in the outcome variables used in the studies, in the treatments implemented, and in the settings in which the studies were conducted. In this review, we have examined information about the characteristics of each study and have attempted to compare and contrast the studies along the dimensions suggested in the Holland and Rubin model and related discussions. Descriptions of the studies and summaries of the results using the approach suggested by Holland and Rubin are given in Section 2. In Section 3, a synthesis of the results and on overall interpretation of the results of the 16 studies is given.

## Section 2: Characteristics and Results of the Studies

As described in Section 1, after a search of the literature and a preliminary review of potential studies, we identified 16 studies which provide evidence related to the educational effectiveness of Sesame Street. The studies varied on a number of dimensions: populations studied, nature of the experimental treatment, outcome variables used as the basis for assessing educational effectiveness, and research design implemented in the study.

To provide an overview of the characteristics of the studies and an opportunity for comparing them with those of other studies, we list the various characteristics in eight tables on the following pages. The studies are compared and contrasted on the following characteristics:

- Experimental and Control Treatments [Tables 1a & 1b];
- Outcome Variables and "Expected Values" of the outcome variables (estimates of what would have happened to the children in the experimental treatment if they had experienced the control treatment instead) [Tables 2a & 2b];
- Populations, Samples, and Concomitant Variables (variables that can be used as covariates or indicators of comparability of experimental and control groups) [Tables 3a & 3b]; and
- Summaries of Results ("effect sizes" given by the researchers or estimated independently for this review; other indicators of significant findings given by the researchers).

Summaries of each study with more detailed descriptions of the results are given in the Appendix.

**Table 1a. Experimental Treatment and Control Treatment (1970-1972)**

| Year | Study | Experimental Treatment | Control Treatment |
|---|---|---|---|
| 1970 | Ball and Bogatz | ▪ Children (self-selected) viewed Sesame Street at home during Year 1 season.<br>▪ Four levels of viewing were determined by researchers based on parent survey.<br>▪ Viewing was not controlled by researchers; some children were encouraged to view. | ▪ Basic Study: No nonviewing comparison group.<br>▪ Age Cohort Study: Children at same age in prior year did not view Sesame Street. No special treatment. |
| 1970 | Miller and Skvarcius; Reeves | ▪ Children (randomly selected) viewed first 60 shows of Year 1 Season -- 5 shows per week for 12 weeks.<br>▪ Viewing controlled by researcher/teacher. | ▪ Randomly selected comparison children played and engaged in other activities. |
| 1971 | Sprigle | ▪ Children (intact groups) viewed Sesame Street in kindergarten during first season.<br>▪ Viewing controlled by researcher/teacher. | ▪ Matched comparison children received a formal language arts program. |
| 1971 | Bogatz and Ball | ▪ Children (randomly selected) viewed Sesame Street at home during Year 2 season.<br>▪ Viewing not controlled by researcher. | ▪ Randomly selected comparison group given no special treatment.<br>▪ Nonviewing controlled by having no broadcast available in children's locality. |
| 1971 | Australian Broadcasting Company | ▪ Sample (random half) of all children (aged 3 to 6 prior to broadcast) in a town in Australia. They were tested in December 1971. Series began in January 1971.<br>▪ Viewing not controlled by researchers. | ▪ Sample (remaining random half) of children (aged 3 to 6 before broadcast). They were tested in December 1970 prior to broadcast of the series. |
| 1972 | Ellis, Reid, & Hoen | ▪ Retrospective study of children in 8 randomly selected kindergarten classes. Two levels of viewing based on parent survey (over 3 year period) compared.<br>▪ Viewing not controlled by researchers. | ▪ No nonviewing comparison group. |
| 1972 | Lukoff | ▪ All children in one Day Care Center viewed 3 Sesame Street shows per week.<br>▪ Viewing controlled by researcher. | ▪ All children in a neighboring center received no special treatment.<br>▪ Nonviewing controlled in center. |
| 1972 | Minton | ▪ Retrospective study of all K children in one school district (1970). Researchers assumed some overall amount of viewing by the group.<br>▪ Viewing not controlled by researcher. | ▪ All kindergarten children in the same district in two prior years (1968, 1969).<br>▪ Nonviewing assured prior to November 1969 (broadcasting began in 1970). |

BEST COPY AVAILABLE

**Table 1b. Experimental Treatment and Comparison Treatment (1972-1987)**

| Year | Study | Experimental Treatment | Control Treatment |
|---|---|---|---|
| 1972 | Salomon, Egistein, Finkelstein, R., Finkelstein, L., Mintzberg, Malve, & Velner | ▪ Children (encouraged/non-encouraged) viewed Sesame Street at home in Year 1 in Israel.<br>▪ Viewing not controlled by researchers. | ▪ No nonviewing comparison group. |
| 1973 | Diaz-Guerrero, Reyes-Lagunes, Witzke, & Holtzman | ▪ Children (randomly selected) in Day Care centers viewed 130 shows of Plaza Sesamo in Year 1 – 5 days a week for 6 months.<br>▪ Viewing controlled by researcher/teachers. | ▪ Randomly selected comparison group viewed cartoons and other non-education films.<br>▪ Nonviewing controlled in day care centers. |
| 1973 | Lemercier & Teasdale | ▪ Retrospective study of 67 kindergarten children (from 9 randomly selected classes). Two levels of viewing based on knowledge of Sesame Street characters.<br>▪ Viewing not controlled by researchers. | ▪ No nonviewing comparison group. |
| 1975 | Darnell & Goodwin | ▪ Retrospective study of 3rd grade children in a random sample of 8 schools (1974). Their viewing in pre-K (1971) had was known.<br>▪ Viewing not controlled by researchers. | ▪ No nonviewing comparison group. |
| 1977 | Taylor & Skanes | ▪ Group 1 (sample of children in locality with Sesame Street available on home TV).<br>▪ Group 2 (sample of children in locality with no TV available at home) viewed Sesame Street 1 hour per day in school on VCR. Viewing controlled by researcher. | ▪ Nonviewing group (randomly selected in a community) of children in locality with no TV. |
| 1977 | Tower, Singer, Singer, & Biggs | ▪ Preschoolers (first of 3 randomly assigned conditions) viewed Sesame Street one-half hour per day for two weeks.<br>▪ Viewing controlled by researchers. | ▪ Comparison group 1 (second of the 3 randomly assigned conditions) viewed Mister Rogers.<br>▪ Comparison group 2 (third of the 3 randomly assigned conditions) viewed Nature/Animal films. |
| 1983 | Owens & Williams | ▪ Group (self-selected) of preschool children (N=15) who watched 5 hours of Sesame Street shows per week (based on parent survey in August 1982).<br>▪ Viewing not controlled by researchers. | ▪ Group (self-selected) of children (N=29) who did not watch Sesame Street at all (based on the same survey). |
| 1987 | Field | ▪ Group of 330 families with 5 year old children. Viewing intensity measures were obtained from diaries kept by families.<br>▪ Viewing not controlled by researchers. | ▪ No nonviewing comparison group. |

**Table 2a. Outcome Variables and Expected Values (1970-1972)**

| Year | Study | Outcome Variables (Y) | Expected Value of ($Y_t$) |
|---|---|---|---|
| 1970 | Ball and Bogatz | ▪ Basic pre-reading skills;<br>▪ Vocabulary;<br>▪ IQ (PPVT). | **Basic Study:**<br>▪ Estimated by the value of the lowest scoring Quartile (Q4), i.e., the estimate for what each of the other Quartiles would have achieved in the absence of Sesame Street is what the lowest viewing quartile scored.<br>**Age Cohort Study:**<br>▪ Expected values of 1970 age cohort estimated by value of corresponding age cohort in 1969 prior to broadcast. |
| 1970 | Miller and Skvarcius; Reeves | ▪ Basic pre-reading skills (Adapted from Ball and Bogatz, 1971). | ▪ Expected value estimated by performance of randomly selected control group. |
| 1971 | Sprigle | ▪ Pre-reading skills;<br>▪ IQ. | ▪ Expected value estimated by performance of comparison group. Initial data used to verify comparability not given. |
| 1971 | Bogatz and Ball | ▪ Basic pre-reading Skills;<br>▪ Vocabulary;<br>▪ IQ. | ▪ Expected value estimated by children in control homes - homes not set up to receive Cable TV system. Pretest data provides evidence for comparability and data for adjusting for initial differences. |
| 1971 | Australian Broadcasting Company | ▪ Vocabulary;<br>▪ Pre-reading skills;<br>▪ IQ. | ▪ Expected value estimated by a randomly selected group of children at the same age in the year prior to broadcast. |
| 1972 | Ellis, Reid, & Hoen | ▪ Knowledge of letters and numbers. | ▪ Expected value estimated by low-viewing children in the original pool. |
| 1972 | Lukoff | ▪ Name, match, and recognize letters, numbers, and shapes. | ▪ Expected value for children in Center 1 estimated by group of children in Center 2. No data given on initial differences.<br>▪ Additional evidence for validity of results given in terms of previous research on letter recognition skills of children when presented in written and oral modes. |
| 1972 | Minton | ▪ Pre-reading skills included on the Metropolitan Readiness Tests (MRT). | ▪ Expected value (for 1970 children) estimated by children in the same kindergarten classes in the preceding years (1969, 1968). |

**Table 2b. Outcome Variables and Expected Values (1972-1987)**

| Year | Study | Outcome Variables (Y) | Expected Value of ($Y_t$) |
|---|---|---|---|
| 1972 | Salomon, Eglstein, Finkelstein, R., Finkelstein, L., Mintzberg, Malve, & Velner | ▪ Basic pre-reading skills;<br>▪ Cognitive styles (field dependence, field independence). | ▪ Estimated using the regression of posttest scores on pretest scores and other non-exposure variables. |
| 1973 | Diaz-Guerrero, Reyes-Lagunes, Witzke, & Holtzman | ▪ Basic pre-reading skills;<br>▪ Field dependence/independence;<br>▪ Oral comprehension;<br>▪ General knowledge. | ▪ Estimated by the performance of randomly selected control group. |
| 1973 | Lemercier & Teasdale | ▪ Basic pre-reading skills;<br>▪ PPVT. | ▪ Expected value for high frequency viewers was the performance of a group of low frequency viewers. |
| 1975 | Darnell & Goodwin | ▪ Basic pre-reading skills – MRT;<br>▪ Reading achievement – Gates-MacGinitie Test. | ▪ No data given. |
| 1977 | Taylor & Skanes | ▪ Basic pre-reading skills;<br>▪ Reading achievement – PPVT, Wechsler Preschool and Primary Scale of Intelligence, Illinois Test of Psycholinguistic Abilities. | ▪ Expected value estimated by the performance of communities with no broadcast of shows available. |
| 1977 | Tower, Singer, Singer, & Biggs | ▪ Recall and recognition of factual information;<br>▪ Inferential skills. | ▪ Expected value estimated by performance of Mister Rogers group. |
| 1983 | Owens & Williams | ▪ Learning styles;<br>▪ Reading readiness skills. | ▪ Expected value estimated by those children not watching Sesame Street. |
| 1987 | Field | ▪ Comprehension;<br>▪ Reflectivity. | ▪ Expected value estimated for high viewers by performance of low viewers. |

**Table 3a. Populations, Samples, and Concomitant Variables (Covariates) (1970-1972)**

| Year | Study | Populations and Samples | Concomitant Variables (Covariates) |
|---|---|---|---|
| 1970 | Ball and Bogatz | ▪ Children between the ages of 3 and 5 in five regions of the United States. (N=943)<br>▪ Two distinct age cohorts (53 to 58 months old) before and after series broadcast. (N1=114; N2=101) | Basic Study:<br>▪ Pretest scores on every child before the start of the broadcast. Positively related to posttest scores and levels of viewing.<br>Age Cohort Study:<br>▪ No covariates in this study. Different children in each year. |
| 1970 | Miller and Skvarcius; Reeves | ▪ Children in 3 Day Care Centers in Maine, New York, and Tennessee. (N=211) | ▪ Pretest scores on every child before the start of the broadcast. Used as a base for gain scores. |
| 1971 | Sprigle | ▪ Children in 2 Head Start Kindergarten programs. (N=48)<br>▪ Children in the same kindergartens in prior years. (N=24) | ▪ Pretest and demographic data used for matching children. |
| 1971 | Bogatz and Ball | ▪ Preschool, urban, disadvantaged children in three cities. (N=283)<br>▪ Two distinct age cohorts (63 to 68 months old) at pretest and posttest. (N=60) | ▪ Pretest data provided evidence for comparability and data for adjusting for initial differences. |
| 1971 | Australian Broadcasting Company | ▪ Children (3-6 years old) in a "fairly large but rather isolated centre of population" in Australia. A random sample of 270 were tested pre broadcast. An independent random sample of 394 were tested post broadcast. | ▪ The authors used demographic variables and other data not related to Sesame Street as evidence of the comparability of the groups. |
| 1972 | Ellis, Reid, & Hoen | ▪ One kindergarten class was randomly selected in each of 8 randomly selected schools in Vancouver. (N=80) | ▪ No data on differences between high and low viewers to alleviate self-selection problem. |
| 1972 | Lukoff | ▪ Children in 2 Day Care centers. (N1=14; N2=25) | ▪ No data given on initial differences.<br>▪ Additional evidence for validity of results given in terms of performance on items less related to the shows -- the groups differed little on those items. |
| 1972 | Minton | ▪ All kindergarten children over a 3 year period in one school district in New York City. (N1=524; N2=495; N3=482) | ▪ Other data were used as evidence for comparability of the groups. |

**Table 3b. Populations, Samples, and Concomitant Variables (Covariates) (1972-1987)**

| Year | Study | Populations and Samples | Concomitant Variables (Covariates) |
|---|---|---|---|
| 1972 | Salomon, Egistein, Finkelstein, R., Finkelstein, I., Mintzberg, Malve, & Velner | ▪ Sample of kindergarten children in Israel. (N=93)<br>▪ Sample of children in Grades 2 and 3. (N=224) | ▪ Pretest scores and demographic variables used in a regression model. |
| 1973 | Diaz-Guerrero, Reyes-Lagunes, Witzke, & Holtzman | ▪ Children in 3 Day Care Centers. (N=173) | ▪ Initial pretest data were collected on all children in the study. |
| 1973 | Lemercier & Teasdale | ▪ Sample of children between the ages of 48 and 63 months in Australia. (N=67) | ▪ Self-selection a problem. No attempt to deal with it. |
| 1975 | Darnell & Goodwin | ▪ A sample of 450 children who started kindergarten in 1971 and finished 3rd Grade (N=122) in 1974. | ▪ Initial data were collected on the children. |
| 1977 | Taylor & Skanes | ▪ A sample of 5, 6, and 7 year old children living in isolated communities on the coast of Labrador. Numbers not given. | ▪ No initial data reported. |
| 1977 | Tower, Singer, Singer, & Biggs | ▪ A sample of preschool children. (N=42) | ▪ Random assignment to control and treatment conditions. No other data beyond the outcome variable collected. |
| 1983 | Owens & Williams | ▪ A sample of primary school children in Mississippi. (N=44) | ▪ No initial data. Self-selection a problem. |
| 1987 | Field | ▪ A sample of 5 year olds and their families. (N=330) | ▪ No initial data. Self-selection a problem. |

## Table 4a. Summary of Results (1970-1972)

| Year | Study | Impact of Sesame Street | Comparison Treatment |
|---|---|---|---|
| 1970 | Ball and Bogatz | ▪ Difference in Gains (Effect Size)<br>Group 1: 19.34 (+0.78)<br>Group 2: 30.80 (+1.14)<br>Group 3: 39.16 (+1.36)<br>Group 4: 48.15 (+1.57)<br><br>▪ [(1970 Cohort) - (1969 Cohort)](Effect Size)<br>Group 1: 11.65 (+0.52)<br>Group 2: 19.73 (+1.04)<br>Group 3: 39.96 (+1.59)<br>Group 4: 40.29 (+2.13) | ▪ Basic Study: No nonviewing comparison group.<br><br>▪ Age Cohort Study: Children at same age in prior year did not view Sesame Street. No special treatment. |
| 1970 | Miller and Skvarcius; Reeves | ▪ $Y_t - Y_c$ (Average Effect Size)<br><br>Total Group: 7.12 (+0.96) | ▪ Randomly selected comparison children played and engaged in other activities. |
| 1971 | Sprigle | ▪ Study 1: (Average Effect Size)<br><br>Total Group: (-3.04)<br><br>▪ Study 2: (Average Effect Size)<br><br>Total Group: (-0.18) | ▪ Matched comparison children received a formal language arts program. |
| 1971 | Bogatz and Ball | ▪ $Y_t - Y_c$ (Average Effect Size)<br><br>▪ Total Group: 15.00 (+0.52)<br><br>▪ [(1971 Cohort) - (1970 Cohort)](Effect Size)<br><br>Total Group: 10.80 (+0.41) | ▪ Randomly selected comparison group given no special treatment. |
| 1971 | Australian Broadcasting Company | ▪ Positive Impact Reported (Data Not Given) | ▪ Randomly selected comparison group in the same town in the year prior to broadcast received no special treatment. |
| 1972 | Ellis, Reid, & Hoen | ▪ Positive Impact Reported (Data Not Given) | ▪ No nonviewing comparison group. |
| 1972 | Lukoff | ▪ Positive Impact Reported (Data Not Given) | ▪ Comparison children in neighboring center received no special treatment. |
| 1972 | Minton | ▪ [1970 Group - 1969 Group](Effect Size)<br><br>▪ Total Group: 1.64 (+0.10)<br><br>▪ [1970 Group - 1968 Group] (Effect Size)<br><br>Total Group: 0.72 (+0.04) | ▪ Comparison kindergarten children in the same district in two prior years (1968, 1969). |

**Table 4b. Summary of Results (1972-1987)**

| Year | Study | Impact of Sesame Street | Comparison Treatment |
|---|---|---|---|
| 1972 | Salomon, Eglstein, Finkelstein, R., Finkelstein, L., Mintzberg, Malve, & Velner | ▪ Study 1: Positive Impact Reported (Effect Size approximately +0.15 to +0.20)<br>▪ Study 2: Positive Impact Reported (Effect size at least +0.27) | ▪ No nonviewing comparison group. |
| 1973 | Diaz-Guerrero, Reyes-Lagunes, Witzke, & Holtzman | ▪ Average Effect Size<br>Total Group: +0.67 | ▪ Randomly selected comparison group viewed cartoons and other non-education films. |
| 1973 | Lemercier & Teasdale | ▪ Average Effect Size<br>▪ Total Group: +0.82 | ▪ No nonviewing comparison group. |
| 1975 | Darnell & Goodwin | ▪ Moderate viewing of Sesame Street was more highly correlated with readiness measures than low or high levels of viewing. (Data Not Given)<br>▪ Viewing uncontrolled by researchers. | ▪ No nonviewing comparison group. |
| 1977 | Taylor & Skanes | ▪ Positive Impact Reported<br>Viewers gained more than nonviewers on relevant measures. | ▪ Nonviewing comparison group of children in locality with no TV. |
| 1977 | Tower, Singer, Singer, & Biggs | ▪ Mixed Impact Reported<br>▪ Average Effect Size: -0.09 | ▪ Comparison group 1 viewed Mister Rogers.<br>▪ Comparison group 2 viewed Nature/Animal films. |
| 1983 | Owens & Williams | ▪ Positive Impact Reported<br>Average Effect Size: +0.23 (Correlation) | ▪ Comparison group of children (N=29) did not watch Sesame Street at all (based on the same survey). |
| 1987 | Field | ▪ Positive Impact Reported<br>Average Effect Size: +0.20 (Correlation) | ▪ No Nonviewing comparison group. |

## Section 3: Synthesis of Research Results

The sixteen research studies reviewed can be classified in a number of ways based on populations of students, nature of the experimental treatment, outcome variables being studied and assessed, and the type of research design implemented in the study. Because the research design is the foundation on which the argument for the effectiveness of an educational program is based, it is used here as a key element for examining the validity of the arguments made by the researchers to justify their conclusions.

In the following pages, the research studies are examined in the three groups below - determined by the nature of the control or comparison group used to estimate the impact or effectiveness of the treatment in each study:

**(1) Studies with No Control Groups:**

- Ball and Bogatz, 1970;
- Ellis, Reid, and Hoen, 1972;
- Salomon, et Al., 1972;
- Lemercier & Teasdale, 1973;
- Darnell & Goodwin, 1975;
- Owens & Williams, 1983; and
- Field, 1987.

**(2) Studies with Nonequivalent Experimental and Control Groups:**

- Ball and Bogatz [Age Cohort Study], 1970;
- Sprigle, 1971;
- Bogatz and Ball [Age Cohort Study], 1971;
- Australian Broadcasting Company, 1971;
- Lukoff, 1972;
- Minton, 1972; and
- Taylor and Skanes, 1977;

**(3) Studies with Random Experimental and Control Groups:**

- Miller and Skvarcius, 1970;
- Bogatz and Ball, 1971;
- Diaz-Guerrero, 1973; and
- Tower, et Al., 1977.

**Studies with No Control Groups**

The first set of studies (Ball and Bogatz, 1970; Ellis, Reid, & Hoen, 1972; Salomon, et al., 1972; Lemercier & Teasdale, 1973; Darnell & Goodwin, 1972; Owens & Williams, 1983; and Field, 1987) included no control groups in their research designs. All of the studies used measures of viewing gathered from the students or their parents and compared children's performance across levels of viewing. Because the children were not assigned to viewing and non-viewing groups prior to their viewing, it is not possible, on the basis of posttest scores alone, to differentiate **posttest differences among the viewing groups due to viewing** from **posttest differences due to differences that existed prior to the children's viewing Sesame Street**. Note that this is true even if the children who viewed the least viewed no Sesame Street programs at all. The fact that some children selected not to view makes it difficult to justify the assumption that there are no differences between them and the children who selected to view. These studies, weakest in terms of research design alone, can be assigned to two categories:

(1) studies in which the researchers gathered posttest data only, i.e., data only after the children in the study had viewed the shows:

- Ellis, Reid, and Hoen, 1972;
- Lemercier & Teasdale, 1973;
- Owens & Williams, 1983;
- Field, 1987; and

(2) studies in which the researchers gathered data before and after the children viewed the shows:

- Ball and Bogatz, 1970;
- Salomon, et Al., 1972; and
- Darnell & Goodwin, 1975.

In discussing these studies, we first consider the studies in which no initial data were collected.

Ellis, Reid, & Hoen (1972) sent a questionnaire to parents of children in 8 kindergarten classes in Vancouver asking them to indicate how often per week (0 through 5 times), on average, their children watched Sesame Street over a three year period. In each class, the investigators selected 5 children who watched 3, 4, or 5 times per week and 5 children who watched 0, 1, or 2 times per week. The children were given a set of tasks related to letters, numerals, and figures. High viewers outperformed low viewers on all tasks. No attention was given to the possibility that the difference might be due to other variables that distinguished the two groups. These results, while positive, were not conclusive.

Lemercier & Teasdale (1973) replicated this study in Australia. They selected 67 children from 9 kindergarten classes and divided them into "high" (N=32) and "low" (N=35) viewers of Sesame Street on the basis of their identification of Sesame Street characters. Minton (1972) had showed that equally large percentages of viewers and nonviewers were able to identify 5 or 6 of the Sesame Street characters, so this procedure for classifying viewers was not credible. Nevertheless, they found that the "high" viewers performed better than the "low" viewers on reading readiness tests. Although positive, these findings are not conclusive.

Owens & Williams (1983) conducted a similar study in the Rocky Mountains. They asked the parents of 44 children how many hours per week the children watched Sesame Street. Those who watched 5 or more hours (N=15) were compared with those (N=29) who did not watch at all. Those who watched outperformed those who did not on a number of the subtests of the Metropolitan Readiness Tests. Although the descriptive results were positive, in the absence of independent information about the comparability of the two groups, the results were not conclusive.

Finally, Field (1987) investigated how often families viewed Sesame Street and other television shows with their five year old children. They found that IQ

scores were positively related to the viewing of Sesame Street. Once again, it was not possible to rule out the effects of other variables that may have distinguished children who viewed more Sesame Street programs from those who did not.

In the three remaining studies, the investigators collected pretest information on the children before viewing took place. In the Ball and Bogatz (1970) study, children were pretested and posttested. In addition, Ball and Bogatz gathered information on demographic characteristics and included children from multiple sites. In the analysis, the children were divided into four quartiles based on reports of viewing from parents. Gains on items and tests in the pretest and posttest batteries were derived for children in each quartile and described. The children in the higher viewing quartiles gained more on average than those in the lower viewing quartiles. However, pretest scores, posttest scores, and gain scores were positively correlated with viewing. Therefore, the viewing groups were not similar on characteristics other than viewing. Therefore, it was not possible in this study to assign a causal effect. Important differences in the characteristics of the children in the viewing groups precluded that. Note that this is true even if the children in the lowest viewing quartile did not view Sesame Street at all. The "age cohort" study in 1970, and the Bogatz and Ball (1971) study in Year 2 of Sesame Street broadcasting, were undertaken to improve the credibility of the first year results and to provide stronger evidence for the positive impact of Sesame Street on children's pre-reading skills.

In the Salomon, et al. (1972) study, researchers pretested children in 1971 before the Sesame Street series was broadcast in Israel. During the season, measures of frequency of viewing were collected. Pretest data, viewing data, and demographic variables were entered into a regression with posttest data as the dependent variable. The investigators reported an impact of Sesame Street viewing independent of pretest scores. This conclusion was based on the fact

that the association between initial ability and viewing measures decreased as the children progressed in school, and, at the same time, the association of viewing measures to posttest increased. Salomon and his colleagues reported that the overall effects of Sesame Street on cognitive skills were greater for children in Grade 2 than for children in kindergarten.

Finally, Darnell and Goodwin (1975) studied 451 children who entered kindergarten in 1971. The researchers followed the children through Grades 1, 2, and 3 and studied the long term effects of Sesame Street viewing on their reading achievement (measured by the Gates-MacGinitie Reading Test) at the ends of Grades 1, 2, and 3. By the end of Grade 3, only 122 of the original 451 children were still in the study. Based on means and standard deviations for demographic variables, the researchers concluded that the initial and final groups of children were comparable. The researchers regressed the Gates-MacGinitie test scores on Metropolitan Reading Test (MRT) scores, teacher judgments of children's reading achievement, age, sex, family SES, preschool experience (Head Start, or other structured experience), and frequency of Sesame Street viewing. They found "very little relationship" between Sesame Street viewing before kindergarten and third grade reading. Given that the researchers had also found a positive relationship between frequency of viewing and MRT scores, the finding in the regression analysis is consistent with results found in other studies.

In discussing their inconclusive results, the authors mention that the MRT test scores for entering kindergarten children increased by half a standard deviation between 1968 and 1972. They also mention that teachers and parents concluded that the "skills and cognitive levels" of entering kindergarten children had been increasing over the period from 1968 to 1975. Although the authors do not conjecture about reasons for those increases, a positive effect due to Sesame Street viewing by all preschoolers in the early years of broadcasting may be a reasonable hypothesis.

**Studies with Nonequivalent Experimental and Control Groups**

There were seven studies in which the researchers used a group that was not randomly selected from the same population as the experimental group to estimate what the performance of the experimental children might have been in the absence of the experimental treatment. Because the groups were not randomly selected, it could not be assumed that the groups were equivalent to the experimental groups.

As in the review of studies with no control groups, these studies can be assigned to categories in terms of whether of not the researchers gathered initial data on the children in the experimental condition:

(1) studies in which the researchers did not gather pretest data on the children in the experimental condition:

- Ball and Bogatz [Age Cohort Study], 1970;
- Bogatz and Ball [Age Cohort Study], 1971;
- Australian Broadcasting Company, 1971;
- Lukoff, 1972;
- Minton, 1972; and

(2) studies in which the researchers gathered pretest and posttest data on the children:

- Sprigle, 1971;
- Taylor and Skanes, 1977;

In discussing these studies, we first consider the studies in which no initial data were collected.

In the Ball and Bogatz [Age Cohort Study] (1970), the researchers used a portion of their sample to fashion a study with a control group. They selected children between the ages of 53 and 58 months at the time of pretesting (1969) to serve as a control (comparison) group children between the ages of 53 and 58 months at the time of posttesting (1970). At posttesting, the children in

the study had viewed Sesame Street to a greater or lesser degree. At pretesting, Sesame Street had not yet begun broadcasting, so none of the children had viewed Sesame Street. Ball and Bogatz compared the posttest data of the 1970 cohort with the pretest data of the 1969 cohort (see Table 4a). By using the 1969 data as an estimate of what the 1970 children would have achieved in the absence of Sesame Street, Ball and Bogatz provided additional evidence for the positive impact of Sesame Street. Of course, an alternative hypothesis that the 1970 cohort was a higher achieving group by chance, or because of some other phenomenon, was not ruled out. Ball and Bogatz found this additional analysis useful and replicated it in the Bogatz and Ball study (1971). The cumulative evidence of Year 1 and Year 2 provided supporting evidence for a positive impact of Sesame Street.

The study conducted by the Australian Broadcasting Company in 1971 provides a replication of the Ball and Bogatz Age Cohorts Study with large samples of children. They compared two randomly selected age cohorts (Ages 3 to 6) before (N=270) and after (N=34) the Sesame Street series was broadcast. The report of their study available at CTW in New York City did not include the numerical results, but the narrative provides a very similar story to that of the Ball and Bogatz study of Year 1. Although the age spread was large, "the average age difference in the two samples was insignificant, less than 1/100th of a month, i.e., less than half a day." "We also found no difference between the "Before" and "After" homes with respect to the amount of television viewed and the amount of A.B.C. or commercial television viewed." Although the investigators had not yet completed their analyses, they reported that, although they found no significant differences on a picture vocabulary (IQ) test, there were significant differences on other tests. Although they did not give the data, they reported only one instance of a negative difference. Other differences were positive, though often not significant. Although the design in this study was a weak one to begin with, the investigators provided additional support that the

1970 data was a reasonable proxy for what the children in 1971 might have achieved had they not been exposed to Sesame Street.

Lukoff (1972) conducted a very small but informative study. She compared children in two Day Care Centers, one of which was known to schedule the viewing of Sesame Street about three times a week with some regularity. She administered very simple exercises to the children - name the letters of the alphabet when shown in groups of four. Then, if the child did not get a perfect score, she asked the child to point to the 'M', the 'D,' and so forth. The same was done with geometric figures and numbers. Based on theories of human development and experience, reading specialists know that children learn some letters sooner than others. Furthermore, recall and recognition skills develop in a corresponding fashion. However, Lukoff noted that data on recall and recognition skills published in the Ball and Bogatz study (1970) showed a greater difference than she would have expected. She thought the unexpected difference might be due to Sesame Street. Lukoff found a similar occurrence in her own data in this small study. Although there was little difference in the performance of the children at the two centers in naming the letters when shown the written symbols, the children in the Sesame Street condition far surpassed the control condition children in pointing out the letters when they were spoken. Lukoff presents this result as potential evidence of a positive impact of Sesame Street.

The study conducted by Minton (1972) was similar to age cohort studies conducted by Ball and Bogatz (1970) and Bogatz and Ball (1971), and the study conducted by the Australian Broadcasting Company (1971). Although Minton did not gather the data herself as the investigators mentioned did, she was able to use archival data (Metropolitan Readiness Test scores) already existing for all the kindergarten children in one school for the years 1968, 1969, and 1970. Minton obtained evidence that all but a tiny (N = 13) fraction of the children in

1970 did in fact view Sesame Street, and a large fraction (54% of the sample) reported that they viewed daily. As in the Australia study, her samples were relatively large (524, 495, and 482). The overall difference on the Metropolitan Readiness Test was small but positive, with only the differences in scores on the Alphabet test in the battery showing a significant positive change. Thus, Minton's research provided support for the impact of Sesame Street on the pre-reading skills measured in the Alphabet subtest of the MRT.

In the remaining two studies, the researchers gathered pretest and posttest information on both groups (Sprigle, 1971; and Taylor and Skanes, 1977). Those data provided the researchers with the opportunity to assess the extent of the differences between the groups before the experimental condition was implemented and to use the pretest data with the posttest data to provide a more credible estimate of the effect of the experimental treatment.

Sprigle conducted two studies of Sesame Street in 1971. Although he did not assign children to treatment and control conditions randomly, Sprigle selected two matched groups (24 pairs) of poverty children from three Head Start kindergarten programs. One group of 24 children viewed Sesame Street programs and worked with related CTW materials in the experimental condition. Children in the control condition were exposed to learning experiences in a game format. Scores on the Metropolitan Readiness Test after three weeks in the first grade showed that the children in the control condition scored higher than the children in the experimental condition. The scores for the control children seemed very high in absolute terms and were reported differently in two publications by Sprigle (1971, 1972). In the first publication, there were surprising IQ differences reported for the matched groups. Those data were not reported in the second publication. Sprigle also compared the Sesame Street group to a group of similar children in the same school in the prior year. Those results showed small differences between the two groups. The control children scored slightly higher on all tests except the Alphabet test. That result is in

keeping with the Minton data and supports the conclusion of a small but positive impact of Sesame Street on the Alphabet subtest of the Metropolitan Readiness Test.

Taylor and Skanes (1977) selected three samples of children from four communities. In community 1, Sesame Street was available on television. In communities 2 and 3, Sesame Street was not available on television, so the researchers could assume reasonably that the children were not exposed to Sesame Street other than incidentally. In community 4, no broadcast was available, but children in kindergarten and first grade viewed Sesame Street for one hour per day on a VCR. Note that, although the researchers selected random samples from each of the communities, the groups are not randomly selected from a common population. They are already divided into pre-existing groups. The researchers gathered pretest data and repeated measures on the children over a three year period using 24 criterion-referenced tests (CRTs).

In this design, the children who have no access to television can provide a reasonable estimate for the performance of the other two groups IF the researchers provide information about the comparability of the groups on the pretest. Unfortunately, they do not. They give only the raw gains of all three groups. We don't know the starting points of the groups. We don't know the relationship between the pretest scores and the posttest scores. If we did, we could make a rather credible estimate of the effectiveness of the VCR treatment and the broadcast television. As it is, the children watching Sesame Street in school may have lost time that would have been spent in a successful language arts program. We can't tell from the data these researchers published. Yet, we know the data were there to conduct a more credible analysis and draw more convincing conclusions.

**Studies with Random Experimental and Control Groups**

Four studies used randomization to establish the experimental and control groups. In three of the studies, the researchers exercised control over the experimental condition which was implemented in a school or center where the children received the treatment as a group. In the fourth study, the children watched Sesame Street at home. In that study, the researchers had no control over the experimental or control conditions.

As in the previous studies, these studies can be divided in terms of the collection of pretest data on the children:

(1) studies in which the researchers did not gather pretest data on the children in the experimental condition:

- Bogatz and Ball, 1971;
- Diaz-Guerrero, 1973; and

(2) studies in which the researchers gathered pretest and posttest data on the children:

- Miller and Skvarcius, 1970;
- Tower, et Al., 1977.

In the first study (Miller and Skvarcius, 1970), children were randomly assigned to the experimental condition (viewed 5 Sesame Street shows per week for 12 weeks) and the control condition (played and engaged in other non-educational activities). No initial pretest data were collected. However, the children remained in their respective groups throughout the course of the study, and there were no indications of differential attrition or other factors that might indicate the loss of the equivalence of the groups. At the end of 12 weeks, the children were tested using a battery of reading readiness tests. The Sesame Street children scored higher than the control children on all tests. A more complete summary of the data is given in the Appendix. An average effect size (+0.96), the difference in posttest scores divided by the standard deviation of

posttest scores in the control group, was calculated and reported in Table 4a. This result provides strong evidence for the conclusion that Sesame Street had a positive impact on the children's pre-reading skills and aptitudes.

In the second study (Diaz-Guerrero, 1973), preschool children in three day care centers were stratified by age and sex and randomly assigned to treatment and control conditions. Children in the treatment condition watched 130 Plaza Sesamo programs five days a week for six months. Children in the control condition watched cartoons. In this study, the children in the Sesame Street condition performed significantly better than the children in the control condition. The complete results are given in the Appendix. An average effect size of +0.67 was calculated for this study. This result provides further evidence for the conclusion that Sesame Street had a positive impact on children's pre-reading skills and aptitudes. It also provides evidence of generalizability to another language and culture.

In the third study (Tower, Singer, Singer, & Biggs, 1977) children were randomly assigned to watch Sesame Street, Mister Rogers, or Nature/Animal films. The investigators focused on the Sesame Street and Mister Rogers groups only. Different measuring instruments (50 questions on a series of ten shows) were created for each group. Questions about the specific shows were scored on memory of factual information, inference, recall, and recognition. The questions were not pretested for level of difficulty. The results (percent of questions answered correctly) indicated that children who watched Sesame Street scored higher on inferential items and recall items. The children who watched Mister Rogers scored higher on factual memory and recognition items. The design in this study was a good one for comparing effects of Sesame Street to effects of Mister Rogers. However, the outcome measures, specific to the individual programs, appear to show that the children learned some things from both shows. As far as the educational effectiveness of Sesame Street is concerned, this study adds neither positive nor negative evidence.

In the fourth study (Bogatz and Ball, 1971), the children were at home and their activities were not monitored. However, the researchers were able to arrange for Sesame Street to be shown over Cable TV to a random sample of the homes in two sites. Children assigned to the control condition did not have access to television. The researchers collected initial pretest data and final posttest data. Although the groups were considered random, and the pretest data confirmed their comparability, Bogatz and Ball used differences in gain scores as their basic measures of effectiveness of Sesame Street. The children who watched the second season of Sesame Street outperformed the control children on all subtests in the test battery. An average effect size of +0.52 was calculated based on the children's total test data. The results provided strong evidence for the researchers' conclusion that Sesame Street had a positive impact on the pre-reading skills and aptitudes of the children.

**Summary and Conclusion**

As in other reviews, the story is a mixed one. There are studies that present results with little credibility. There are other studies that present negative results - the most serious one being that of Sprigle (1971). There are problems that can be raised to cast doubt on the results in the Sprigle study. There is the possibility that other studies with negative findings were not published because of a reluctance to report negative findings.

On the whole, however, the studies with the strongest designs indicate that Sesame Street had a significant positive impact on the children in terms of the variables measured and relative to the children in the other groups studied. And, the magnitude of the impact was large enough to have shown up in studies with very small samples. Therefore, based on the research studies reviewed in this report, Sesame Street, over the past twenty years, has had a significant positive impact on the pre-reading and school-readiness skills of children in the United States, and of children in at least four other countries - Australia, Canada, Israel, and Mexico.

## REFERENCES


Australian Broadcasting Corporation. (1971). SESAME STREET: An ABC research interim report. Sydney: Author.

Ball, S. & Bogatz, G. A. (1970). The first year of SESAME STREET: An evaluation. Princeton, NJ: Educational Testing Service.

Bogatz, G. A. & Ball, S. (1971). The second year of SESAME STREET: A continuing evaluation. (Vols. 1 & 2). Princeton, NJ: Educational Testing Service. (ERIC Document Reproduction Service Nos. ED 122 800, ED 122 801).

Bogatz, G. A. & Ball, S. (1971). A summary of the major findings in "The Second Year of SESAME STREET: A Continuing Evaluation." Princeton, NJ: Educational Testing Service. (ERIC Document Reproduction Service No. ED 122 802).

Darnell, C. D. & Goodwin, W. L. (1975, August). The kindergarten child 1971 or the class of 1984. Paper presented at the 83rd Annual Meeting of the American Psychological Association, Chicago, Illinois. (ERIC Document Reproduction Service No. ED 116 816).

Diaz-Guerrero, R. (1973). An evaluation of the first season of PLAZA SESAMO, an educational television program for preschool children in Mexico. Mexico City: National University of Mexico, Centro de Investigaciones Psicopedagocicas.

Diaz-Guerrero, R. & Holtzman, W. H. (1974). Learning from televised PLAZA SESAMO. Journal of Educational Psychology, 66(5), 632-643.

Ellis, E. N., Reid, M. J. & Hoen, R. (1972). The impact of SESAME STREET on primary pupils in Vancouver (RR-72-16). Vancouver, British Columbia: Department of Planning and Evaluation. (ERIC Document Reproduction Service No. ED 077 988).

Field, D. E. (1987). Child and parent coviewing of television: Its extent and its relationship to cognitive performance (Doctoral Dissertation, University of Massachusetts). Dissertation Abstracts International, 48(09), 2799B.

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. Educational Researcher, 5, 3-8.

Glass, G. V. & Smith, M. L. (1976). Meta-analysis of psychotherapy outcome studies. Paper presented at the Annual Meeting of the Society for Psychotherapy Research, San Diego, CA.

Holland, P. W. (1986). Which comes first, cause or effect? The New York Statistician, 38, 1-6.

Holland, P. W. & Rubin, D. B. (1983). On Lord's paradox, in H. Wainer and S. Messick (eds.) Principals of Modern Psychological Measurement. Hillsdale, NJ: Lawrence Erlbaum Associates.

Holland, P. W. & Rubin, D. B. (1987). Causal inference in retrospective studies. Princeton, NJ: Educational Testing Service Program Statistics Research Technical Report No. 87-73.

Kulik, C. C., Kulik, J. A., & Cohen, P. A. (1980). Effectiveness of computer-based teaching: A meta-analysis of findings. Review of Educational Research, 50(4), 525-544.

Lemercier, K. I. & Teasdale, G. R. (1973). SESAME STREET; Some effects of a television programme on the cognitive skills of young children from lower SES backgrounds. Australian Psychologist, 8(1), 47-51.

Lesser, G. S. (1974). Children and television: Lessons from SESAME STREET. New York, NY: Random House, Inc.

Light, R. J. and Pillemer, D. B. (1984). Summing Up: The Science of Reviewing Research. Cambridge, MA: Harvard University Press.

Lukoff, J. (1972). Recall and recognition of letters and numbers of children in two day care centers. New York, NY: Children's Television Workshop.

Messick, S. J. (1980). The effectiveness of coaching for the SAT: Review and reanalysis of research from the fifties to the FTC. Princeton, NJ: Educational Testing Service.

Miller, J. & Skvarcius, R. (1970). Does SESAME STREET teach? New York, NY: Children's Television Workshop.

Minton, J. H. (1972). The impact of SESAME STREET on reading readiness of kindergarten children (Doctoral dissertation, Fordham University). Dissertation Abstracts International, 33(07), 3396.

Owens, J. & Williams, C. C. (1983, November). The effects of viewing SESAME STREET on learning styles and readiness of primary children. Paper presented at the Meeting of the Mid-South Educational Research Association, Nashville, TN.

Reeves, B. F. (1970). The first year of SESAME STREET: The formative research. Final report (Vol. II). New York, NY: Children's Television Workshop.

Rosenthal, R. (1984). Meta-analysis procedures for social research. Applied Social Research Methods Series, Volume 6. Beverly Hills, CA: Sage publications.

Rubin, D. W. (1974). Assignment of treatment group on the basis of a covariate. Journal of Educational Statistics, 6, 1-26.

Salomon, G. (1974). SESAME STREET in Israel: Its instructional and psychological effects on children. Jerusalem, Israel: Hebrew University of Jerusalem.

Salomon, G., Eglstein, S., Finkelstein, R., Finkelstein, I., Mintzberg, E., Malve, D., & Velner, L. (1972). Educational effects of SESAME STREET on Israeli children (brief summary). Jerusalem, Israel: Hebrew University of Jerusalem. (ERIC Document Reproduction Service No. ED 070 317)

Slavin, R. E. (1986). Best-evidence synthesis: An alternative to meta-analytic and traditional reviews. Educational Researcher, 15(9), 5-11.

Slavin, R. E. & Madden, N. A. (1989). What works for students at risk: A research synthesis. Educational Leadership, 4-13.

Sprigle, H. A. (1971). Can Poverty Children live on SESAME STREET? Young Children, 26(4), 202-217. (ERIC Document Reproduction Service No. EJ 034 509).

Sprigle, H. A. (1972). Who wants to live on SESAME STREET? Young Children, 28(2), 91-109. (ERIC Document Reproduction Service No. EJ 071 399).

Taylor, L. J. & Skanes, G. R. (1977). The effects of SESAME STREET in isolated communities. Paper presented at the Annual Meeting of the American Educational Research Association, New York, NY. (ERIC Document Reproduction Service No. ED 136 980).

Tower, R. B., Singer, D. G., Singer, J. L., & Biggs, A. (1977). Differential effects of television programming on preschoolers' cognition, imagination and social play. Connecticut: Yale University and the University of Bridgeport. (ERIC Document Reproduction Service No. ED 153 713).

Tower, R. B., Singer, D. G., Singer, J. L., & Biggs, A. (1979). Differential effects of television programming on preschoolers' cognition, imagination, and social play. American Journal of Orthopsychiatry, 49(2), 265-281.

# APPENDIX

## SUMMARIES OF RESEARCH STUDIES

## SUMMARIES OF RESEARCH STUDIES

In this appendix, studies are listed in chronological order and described briefly. In the summer and fall of 1989, more than 200 potential studies were reviewed at the Children's Television Workshop's New York City offices. Although all of the references reviewed were related to Sesame Street and education, fewer than twenty studies contained data on student performance on educational variables, such as, prereading skills, reading or language skills, or reading comprehension. In some cases, several authors published reports on the same data for different purposes. For example, Miller and Skvarcius (1970) provided a summative report on the same data used by Reeves (1970) for a formative study. In other cases, additional analyses of the same data were presented in later publications. For example, the data in the study of the Sesame Street adaptation in Mexico were presented somewhat differently in 1973 and 1976.

Each study is discussed in two parts: **(1) Treatment and Design** and **(2) Analysis and Results**. Although treatment and design are distinct, authors generally discussed them together. The outcome or dependent variables are discussed in the Analysis and Results section. Frequently, authors do not distinguish between variables, tests, performance tasks, and the measures assigned to the variables. In this report, a similar level of precision has been used in treating those elements of the studies.

## 1. Ball and Bogatz (October 1970)

Ball and Bogatz investigated the effectiveness of Sesame Street in its first full season. They focused on children between the ages of 3 and 5 in five geographical regions of the United States.

**Treatment and Design**

Ball and Bogatz had no control over the treatment at first. They selected a sample of more than one thousand preschool children in five sites: Boston, Suburban Philadelphia, Durham, Phoenix, and Northeastern Rural California. Sites were chosen to satisfy the following criteria:

- Sesame Street was broadcast on VHF daily at 9 a.m. or 10 a.m., and
- The sites included large numbers of disadvantaged children.

The treatment was the extent of "exposure to" or "viewing by" the children. In a sense, all children in the study were in a treatment group with four levels of viewing frequency - four quartiles ranging from lowest frequency to highest frequency of viewing. The authors provided additional encouragement to view the Sesame Street programs to half the children in the study.

The design was a complex one with a variety of settings (At home, In nursery School, or In kindergarten; Encouraged to view or Not encouraged to view; Observed or Not observed; Site 1, 2, 3, 4, or 5; Spanish speaking or Not Spanish speaking). In general, this was a weak design in that all children self-selected themselves into the viewing groups. The researchers had no control over viewing. This is frequently the case in studies of educational programs. If selection is related to any important variable, for example, motivation, it becomes very difficult for the researcher to untangle the causes of any effect found in the study.

Holland and Rubin discuss this difficult issue and suggest that the point be made as explicit as possible using direct and indirect causal terminology. In this study, the authors are aware of the dilemma and explain their view that encouragement and exposure are both indirect causes of the children's learning. By making the definition of "effect" or "impact" explicit, Holland and Rubin are able to clarify, though not solve, the issues involved. The difficulty is in determining an estimate of how the child would have performed had he or she not been in the treatment condition. The "effect" is defined by Holland and Rubin as the difference between the child's performance in the treatment condition and the same child's "hypothetical" performance if he or she had not been in the treatment condition, i.e.,

**Effect = [Observed Performance] - [Estimated Performance].**

The challenge to the researcher is to provide a credible estimate of the child's performance when not in the treatment. Because any individual child may have very unique characteristics, estimates are ordinarily made in terms of the average across many children in the treatment. In most cases, the estimate is that of the performance of similar children in a similar situation that differs from the treatment situation in only one way - absence of the treatment itself. Because that is strictly impossible (something must take the place of missing treatment), researchers attempt to provide a variety of estimates across populations and situations to rule out other plausible alternatives for the cause of the effect. The design is the means by which the researcher tries to insure that a conclusion will follow validly from the analysis of the data based on the design.

In many small research studies, with only two classes of children and a treatment that includes instruction by the teacher, the teacher as a competing plausible cause is never ruled out. Because we know that teachers differ in their effectiveness, two group studies with different teachers are virtually useless in

research. Although there are some such studies in the Sesame Street research, arguments are made that teachers played little or no role in the treatment.

This discussion will be useful in judging the degree of confidence that can be placed in the results of this research study and the other research studies in this review. In order to estimate the performance of T in the absence of the treatment, the researchers suggest using the performance of those who viewed less as an estimate of the performance of those who viewed more. Once all children had the opportunity to watch Sesame Street, that approach meant that Student X1's performance in the absence of the treatment was estimated by Student X2 who did not watch Sesame Street. Since the basic idea is that Student X2 should differ in no way from Student X1, other than being in the treatment group, there is a problem in this design. Student X2 elected not to watch Sesame Street when it was available - perhaps an indication of low motivation.

The authors recognized this weakness and provided an additional comparison group in a special study which they call the "Age Cohort Study." Although this study is a longitudinal one in the sense that the same children are pretested and posttested, it is fortunate in that children at ages 3, 4, and 5 were included in the study. At the time of posttesting, the entire sample of children had grown older. The design therefore provides a way of comparing a group of children in the treatment with a group of children at exactly the same age before Sesame Street went on the air - a type of benchmark measure. The results of that study are given in the tables under Study 2 on the following pages.

**Analysis and Results**

The following dependent or outcome variables were used in this study to assess Sesame Street's educational effectiveness:

| | |
|---|---|
| 1. Body Parts - | Recognize body parts and their functions.<br>Name body parts and their functions. |
| 2. Letters - | Recognize letters (when seen and heard).<br>Name letters.<br>Match letters in words.<br>Recognize letters in words.<br>Initial sounds.<br>Reading words. |
| 3. Forms - | Recognize forms.<br>Name forms. |
| 4. Numbers - | Recognize numbers.<br>Name numbers.<br>Numerosity.<br>Counting.<br>Addition and subtraction. |
| 5. Relational Terms - | Amount relations.<br>Size relations.<br>Position relations. |
| 6. Sorting Skills | |
| 7. Classification - | By size.<br>By form.<br>By number.<br>By function. |
| 8. Puzzles. | |

The authors used viewing data to divide the children into four quartiles based on extent of viewing. The results are given on the following pages. The results are straightforward. Those children who viewed more did better on the tests and exercises administered. In the tables under Study 1, the results are presented in terms of the pretest scores of the children and the gains made by the time of the

posttest. I have computed the "effect size" for each variable, which are sizable of course since the children are older. The important distinction is that the size of the gain increases with increasing amounts of viewing.

The authors realized that the weakness of the design made it difficult to ascribe the increased performance to viewing and to rule out the competing possibility that the more motivated children may have both selected to watch Sesame Street more and performed better on the tests. It was difficult to make the case that the performance of a low viewing child provided a valid estimate of the performance of a high viewing child. Therefore, the researchers suggested that additional evidence could be provided by comparing the posttest performance of the 4 year olds with the pretest performance of the children at the same age before the series was broadcast. The results are given in the tables under Study 2.

This original evaluation study of the first year of Sesame Street provided a convincing argument for the effectiveness of Sesame Street. The authors demonstrated that the results were similar across sex, age, and sites. Those details are not included in the research synthesis at present.

**Study 1: Performance Gain by Viewing Frequency**

**Group: Q1 - Lowest Quartile of Viewers**

| Variables | Pretest N=231 | Gain | SD | Effect Size |
|---|---|---|---|---|
| Body Parts | 18.13 | 3.98 | 6.86 | 0.58 |
| Letters | 13.24 | 4.44 | 6.04 | 0.74 |
| Forms | 8.55 | 2.39 | 3.47 | 0.69 |
| Numbers | 16.38 | 5.69 | 8.31 | 0.68 |
| Matching | 8.01 | 1.20 | 2.65 | 0.45 |
| Relations | 9.11 | 1.19 | 2.96 | 0.40 |
| Sorting | 2.29 | 0.50 | 1.31 | 0.38 |
| Classification | 10.65 | 1.82 | 4.05 | 0.45 |
| Puzzles | 1.93 | 0.45 | 1.37 | 0.33 |
| GRAND TOTAL | 76.32 | 19.34 | 24.73 | 0.78 |
| PPVT Raw Score | 32.26 | NA | 11.27 | NA |

**Group: Q2 - Second Lowest Quartile of Viewers**

| Variables | Pretest N=242 | Gain | SD | Effect Size |
|---|---|---|---|---|
| Body Parts | 20.60 | 4.24 | 6.44 | 0.66 |
| Letters | 14.54 | 8.79 | 7.23 | 1.22 |
| Forms | 10.09 | 3.43 | 3.91 | 0.88 |
| Numbers | 19.23 | 9.14 | 9.30 | 0.98 |
| Matching | 8.55 | 1.40 | 2.48 | 0.56 |
| Relations | 9.93 | 1.64 | 2.93 | 0.56 |
| Sorting | 2.58 | 0.90 | 1.40 | 0.64 |
| Classification | 12.14 | 3.38 | 4.48 | 0.75 |
| Puzzles | 2.10 | 0.82 | 1.33 | 0.62 |
| GRAND TOTAL | 86.29 | 30.80 | 27.11 | 1.14 |
| PPVT Raw Score | 35.90 | NA | 12.18 | NA |

Study 1: (continued)

Group: Q3 - Second Highest Quartile of Viewers

| Variables | Pretest N=235 | Gain | SD | Effect Size |
|---|---|---|---|---|
| Body Parts | 22.27 | 4.29 | 6.54 | 0.66 |
| Letters | 16.16 | 13.51 | 7.98 | 1.69 |
| Forms | 10.65 | 4.18 | 3.61 | 1.16 |
| Numbers | 21.92 | 11.25 | 10.54 | 1.07 |
| Matching | 9.13 | 0.96 | 1.97 | 0.49 |
| Relations | 10.43 | 1.67 | 2.65 | 0.63 |
| Sorting | 2.69 | 1.39 | 1.50 | 0.93 |
| Classification | 12.86 | 4.60 | 4.71 | 0.98 |
| Puzzles | 2.32 | 0.87 | 1.37 | 0.64 |
| GRAND TOTAL | 94.10 | 39.16 | 28.70 | 1.36 |
| PPVT Raw Score | 37.71 | NA | 13.39 | NA |

Group: Q4 - Highest Quartile of Viewers

| Variables | Pretest N=235 | Gain | SD | Effect Size |
|---|---|---|---|---|
| Body Parts | 23.41 | 4.73 | 5.82 | 0.81 |
| Letters | 17.88 | 17.60 | 9.72 | 1.81 |
| Forms | 11.11 | 5.34 | 3.48 | 1.53 |
| Numbers | 24.62 | 13.34 | 11.04 | 1.21 |
| Matching | 9.34 | 1.02 | 1.65 | 0.62 |
| Relations | 10.55 | 2.28 | 3.04 | 0.75 |
| Sorting | 2.77 | 1.70 | 1.37 | 1.24 |
| Classification | 13.51 | 5.26 | 4.63 | 1.14 |
| Puzzles | 2.65 | 0.83 | 1.41 | 0.59 |
| GRAND TOTAL | 100.94 | 48.15 | 30.62 | 1.57 |
| PPVT Raw Score | 41.00 | NA | 11.95 | NA |

Study 2: Age Cohorts Study (Year 1)

Group: Q1 - Lowest Quartile of Viewers

| Variables | Treatment N=26 | Control N=31 | SD | Effect Size |
|---|---|---|---|---|
| Body Parts | 21.04 | 17.87 | 6.49 | 0.49 |
| Letters | 14.65 | 14.06 | 6.45 | 0.09 |
| Forms | 11.04 | 7.45 | 3.36 | 1.07 |
| Numbers | 19.00 | 16.77 | 7.06 | 0.32 |
| Matching | 9.31 | 7.97 | 2.93 | 0.46 |
| Relations | 10.65 | 9.16 | 2.35 | 0.63 |
| Sorting | 2.69 | 2.13 | 1.38 | 0.41 |
| Classification | 11.96 | 10.71 | 3.84 | 0.33 |
| Puzzles | 2.31 | 2.03 | 1.56 | 0.18 |
| GRAND TOTAL (20 | 88.42 | 76.77 | 22.27 | 0.52 |
| PPVT IQ | 81.08 | 75.97 | 26.63 | 0.19 |

Group: Q2 - Second Lowest Quartile of Viewers

| Variables | Treatment N=33 | Control N=33 | SD | Effect Size |
|---|---|---|---|---|
| Body Parts | 22.91 | 20.24 | 5.74 | 0.47 |
| Letters | 18.24 | 13.09 | 3.65 | 1.41 |
| Forms | 11.21 | 9.09 | 3.21 | 0.66 |
| Numbers | 23.76 | 17.97 | 7.10 | 0.82 |
| Matching | 9.97 | 8.45 | 1.99 | 0.76 |
| Relations | 11.30 | 10.33 | 2.98 | 0.33 |
| Sorting | 3.33 | 1.67 | 1.29 | 1.29 |
| Classification | 13.79 | 11.03 | 2.91 | 0.95 |
| Puzzles | 2.55 | 2.55 | 1.37 | 0.00 |
| GRAND TOTAL (20 | 101.70 | 81.97 | 18.90 | 1.04 |
| PPVT IQ | 85.09 | 80.03 | 26.63 | 0.19 |

Study 2: (continued)

Group: Q3 - Second Highest Quartile of Viewers

| Variables | Treatment N=18 | Control N=27 | SD | Effect Size |
|---|---|---|---|---|
| Body Parts | 26.83 | 21.93 | 5.57 | 0.88 |
| Letters | 26.83 | 14.81 | 5.90 | 2.04 |
| Forms | 14.22 | 9.93 | 4.08 | 1.05 |
| Numbers | 32.67 | 20.37 | 9.42 | 1.31 |
| Matching | 10.33 | 8.78 | 2.28 | 0.68 |
| Relations | 12.39 | 10.81 | 2.32 | 0.68 |
| Sorting | 4.28 | 2.81 | 1.55 | 0.95 |
| Classification | 17.78 | 12.89 | 4.50 | 1.09 |
| Puzzles | 3.44 | 2.26 | 1.02 | 1.16 |
| GRAND TOTAL (20 | 130.33 | 90.37 | 25.21 | 1.59 |
| PPVT IQ | 88.33 | 82.67 | 19.28 | 0.29 |

Group: Q4 - Highest Quartile of Viewers

| Variables | Treatment N=24 | Control N=23 | SD | Effect Size |
|---|---|---|---|---|
| Body Parts | 26.75 | 22.87 | 5.74 | 0.68 |
| Letters | 31.92 | 18.52 | 3.65 | 3.67 |
| Forms | 15.46 | 10.35 | 3.21 | 1.59 |
| Numbers | 35.54 | 23.96 | 7.10 | 1.63 |
| Matching | 10.00 | 9.17 | 1.99 | 0.42 |
| Relations | 18.00 | 10.26 | 2.98 | 2.60 |
| Sorting | 4.54 | 2.30 | 1.29 | 1.74 |
| Classification | 17.75 | 13.04 | 2.91 | 1.62 |
| Puzzles | 2.92 | 2.52 | 1.37 | 0.29 |
| GRAND TOTAL (20 | 139.33 | 99.04 | 18.90 | 2.13 |
| PPVT IQ | 88.08 | 86.61 | 26.63 | 0.06 |

## 2. Miller and Skvarcius (October 1970)
## Reeves (December 1970)

As part of its formative research and evaluation, Children's Television Workshop directed Barbara Reeves to set up a formative type study of the initial Sesame Street programs as the series began broadcasting in November 1969 in order to provide feedback to the program producers for use in their continuing development and production of new programs in the series. Therefore, shortly before the series went on the air, three Day Care Centers (Maine, New York, and Tennessee) were asked to collaborate in a study of the early Sesame Street programs. Reeves focused on how the components of the programs (format, pace, segments, and so forth) held the attention and interest of the children. In addition, children were tested at several points during the first six months of programming. The test data were used by Miller and Skvarcius to investigate the program's success in attaining a number of its stated goals. Miller and Skvarcius, as well as Reeves, published the results of this early summative type investigation.

**Treatment and Design**

Half the children in each site were randomly selected to watch the programs while the nonviewers played outdoors or engaged in other activities. An extensive set of tests was administered individually to all children in both groups before the series began and at the end of three months.

The treatment condition was the viewing of Sesame Street programs each day, five days per week, for 12 weeks. Adults were present while the children viewed the programs. The adults were asked not to provide follow up or review. Instructors were asked not to provide deliberately related instruction. No

attempt was made to eliminate or modify any part of the normal day-care instructional program.

In terms of the Holland and Rubin approach, the randomly selected nonviewers served as a control group and their performance was used to estimate how the viewers would have performed in the absence of the Sesame Street program. The pretest and posttest tests were the same.

Although the children were randomly assigned, the pretest data provided a way of checking the success of the random assignment. In addition the authors examined race, IQ, sex, and age to verify and validate the equivalence of the groups. These data also provided the capability to verify findings in a number of populations.

Using the Holland and Rubin approach to estimating causal effects of educational programs, we assume that a process was used to randomly select the children such that each child was equally likely to be chosen for the Sesame Street group. Although this sounds simple, there are many factors in which assignment can be biased. Therefore, researchers generally pretest anyway to verify the similarity of the groups and the "success" of the random assignment. In this case, none of the subtest scores differed significantly across groups. On the PPVT, for example, viewers had a mean score of 97.38; nonviewers had a score of 98.24.

**Analysis and Results**

At the end of three months, the researchers found that the children in the control group were unable to correctly identify members of the Sesame Street cast. However, by the end of six months, the control children were able to identify the Sesame Street characters.

The pretest scores were not precisely equal. In fact, the treatment group had a mean below the control group. We didn't expect them to match perfectly. The difference between the means was less than a quarter of a standard deviation... that means that many of the children in one group outperformed many of the children in the other group. They had a significant overlap. If that were not the case, we would be very concerned about the adequacy of the design.

The variables are listed in the table showing the posttest scores of the two groups. Since the initial groups were randomly selected, we estimated an "effect size" relatively easily using the posttest scores. However, we knew the children in the treatment group were poorer to start with. One way to approach the problem was to compare the simple gain scores. The treatment group gained more than the control group. Yes, but since they started lower, they had more room in which to gain. How about relative gain in terms of the maximum achievable score? But we know that higher achieving children score even higher in the gain score sense - the spread between lower and higher achieving scorers spreads out as both gain? A very standard way to approach this design problem, especially in educational studies, is to assess the dependency of the dependent variable on one or more independent variables. Research studies show that the dependent variable is frequently related to initial aptitude and/or achievement scores.

"MILLER and SKVARCIUS" - October 1970
"REEVES" - December 1970

Group 1: Children in 3 Day Care Centers

| Variables | Treatment N=110 | Control N=101 | SD | Effect Size |
|---|---|---|---|---|
| Body Parts | 38.02 | 37.48 | 0.87 | 0.62 |
| Letters | 38.84 | 35.41 | 2.26 | 1.52 |
| Forms | 5.59 | 4.92 | 0.37 | 1.81 |
| Numbers | 35.16 | 34.61 | 2.09 | 0.26 |
| Sorting Skills | 10.63 | 9.90 | 0.38 | 1.92 |
| Relationships | 8.35 | 8.46 | 0.31 | -0.35 |
| Classifications | 13.93 | 12.99 | NA | NA |
| Puzzles | 7.58 | 7.19 | 0.41 | 0.95 |
| Composite | 158.09 | 150.97 | NA | NA |

## 3. Sprigle (March 1971)

Herbert Sprigle also arranged for two groups of children to serve as treatment and comparison children in a small study during Sesame Street's first broadcast year. The treatment children attended two Head Start kindergarten programs in which Sesame Street and related materials distributed by CTW were used as the basis for an educational component. The comparison children attended another school program in which they were exposed to, in Sprigle's words, "numbers, letters, shapes, spatial and temporal relationships, language, communication, and listening experiences."

**Treatment and Design**

Sprigle selected 24 pairs of children matched on a number of variables that are known to be related to levels of achievement in school (Binet IQ scores, age, parent education, parent occupation, and family income).

To determine the impact of Sesame Street on the viewing children, Sprigle compared their performance on a number of tests to the children in the regular program. In the Holland and Rubin model, his design was a reasonable one. If the children were not watching the program, then they would be participating in their regular kindergarten instructional program - a program which Sprigle hypothesized was successful in enhancing the learning of the children. In fact, his results supported his hypothesis. Mabel Rice (1989) has presented additional evidence that teachers and parents are reluctant to allow Sesame Street programs to replace ordinary learning activities in kindergarten and nursery school programs.

**Analysis and Results**

The results are shown in the tables (Study 1) on the next page. In general the differences on the variables listed are quite large. However, the differences in the Mean IQ scores can be seen to be very large also. Therefore, it is plausible that some of the differences in the outcome measures were due to initial differences between the groups. Sprigle, realizing that the results of Study 1 did not eliminate that possibility, also compared the treatment children to a group of children in the same Head Start programs the year before. The results of that study are shown in the Study 2 table. Although the results still show a positive result for the "regular" kindergarten instructional program, the differences are much smaller.

In general, this study provides evidence that school-based programs like those in the Sprigle study should not be simply replaced by a set of Sesame Street programs without some specific reason or plan in mind. As Rice (1989), mentioned above, has shown, it might be possible to improve overall instruction by the use of programs targeted at specific skills. Further research is needed to provide evidence on that issue.

"SPRIGLE" - March 1971 and December 1972

Study 1: Children in Head Start Kindergarten Programs

| Variables | Treatment N=24 | Control N=24 | T-Value | Effect Size |
|---|---|---|---|---|
| Word Meaning | 4.76 | 7.75 | 5.54 | -1.60 |
| Listening | 6.38 | 11.04 | 7.31 | -2.11 |
| Matching | 5.57 | 10.08 | 4.56 | -1.32 |
| Alphabet | 6.76 | 15.70 | 10.37 | -2.99 |
| Numbers | 7.42 | 16.08 | 10.39 | -3.00 |
| Copying | 3.42 | 9.70 | 7.80 | -2.25 |
| Total | 34.38 | 70.37 | 10.54 | -3.04 |
| Mean IQ (Boys) | 84.00 | 120.00 | | |
| Mean IQ (Girls) | 82.00 | 110.00 | | |

Study 2: Comparison of Sesame Street Viewers with Prior Year's Children

| Variables | Treatment N=24 | Control N=24 | T-Value | Effect Size |
|---|---|---|---|---|
| Word Meaning | 4.76 | 5.45 | 1.11 | -0.32 |
| Listening | 6.38 | 7.45 | 1.32 | -0.38 |
| Matching | 5.57 | 5.80 | 0.19 | -0.05 |
| Alphabet | 6.76 | 6.30 | 0.44 | -0.13 |
| Numbers | 7.42 | 8.00 | 0.57 | -0.16 |
| Copying | 3.42 | 3.90 | 0.42 | -0.12 |
| Total | 34.38 | 36.70 | 0.61 | -0.18 |
| Mean IQ (Boys) | 84.00 | NA | NA | NA |
| Mean IQ (Girls) | 82.00 | NA | NA | NA |

## 4. Bogatz and Ball (November 1971)

In the second year, Bogatz and Ball were able to replicate their findings, and provide information on additional variables, using a design stronger than the one used in the year 1 study. In the second year, sites where Sesame Street was not available were identified. A randomly selected group was then able to view Sesame Street on Cable TV only.

**Treatment and Design**

It is important to understand that this second study is not a replication of the first Ball and Bogatz (1970) study. A new design was implemented in an attempt to improve the validity of the study so that conclusions could be drawn with greater confidence. In addition, an entirely new set of Sesame Street programs were shown in the second season. This was not a second attempt to validate the success of the first season's shows on a new set of children. Both the treatment and the design changed. However, the treatment is considered as one instance of a full complement of Sesame Street programs.

In the second year, the goals targeted by the producers expanded from 40 in Year 1 to 63 in Year 2. In general, the initial goals remained and new goals were added. That became the general paradigm for the next 18 years. While continuing to target the pre-reading and early language skills needed by children, CTW added more complex cognitive skills and attitudes to their production goals. This increase in the number of goals required that the ETS instruments be revised to measure the program's impact in as many goal areas as possible.

In implementing a new design, a sample of 283 urban, disadvantaged children who had not seen Sesame Street in Year 1 (because Sesame Street was

either not broadcast at all in the area or was only broadcast on UHF) were identified in two cities - one in the East and one in the West. With the cooperation of a new Cable TV system in one city, arrangements were made to control which homes in the test community received the shows and which did not. Of course, some children eventually viewed some shows at the homes of friends or in other ways. In the second site, Sesame Street was shown on UHF only. The researchers again used an encouraged/not-encouraged approach to augment the design.

**Analysis and Results**

The results are shown in the tables on the following pages. As in Year 1, the children who watched Sesame Street programs performed better than those who did not on the educational variables included in the study. Bogatz and Ball also analyzed the results for an "age cohort" of children as they had done in the first year study. Children who were between the ages of 63 months and 68 months at the time of the posttests were compared to children who were between the ages of 63 months and 68 months at the time of the pretests. Those results supported and reinforced the overall positive results of the second year study by Bogatz and Ball.

"BOGATZ and BALL" - November 1971

Study 1: Encouraged/Viewers vs Not-encouraged/Non-viewers

| Variables | Treatment N=130 | Control N=153 | SD | Effect Size |
|---|---|---|---|---|
| Body Parts | 14.10 | 12.80 | 4.90 | 0.27 |
| Alphabet (A - Z) | 11.10 | 8.70 | 5.90 | 0.41 |
| Forms | 4.40 | 3.40 | 2.40 | 0.42 |
| Numbers (1 - 30) | 13.20 | 10.90 | 7.00 | 0.33 |
| Numbers Total | 22.60 | 18.50 | 7.60 | 0.54 |
| Relations | 11.50 | 10.20 | 3.70 | 0.35 |
| Sorting | 5.80 | 4.40 | 3.00 | 0.47 |
| Classification | 8.30 | 7.10 | 4.70 | 0.26 |
| Parts of Whole | 4.80 | 4.60 | 1.80 | 0.11 |
| Grand Total | 98.50 | 83.50 | 28.70 | 0.52 |
| PPVT Raw Score | 31.70 | 28.90 | 12.70 | 0.22 |
| PPVT Mental Age | 40.60 | 37.90 | 13.20 | 0.20 |
| Chronological Age | 50.40 | 49.40 | 9.70 | 0.10 |

Study 2: Age Cohorts Study (Year 2)

| Variables | Treatment N=31 | Control N=29 | SD | Effect Size |
|---|---|---|---|---|
| Body Parts | 16.10 | 15.60 | 2.70 | 0.19 |
| Alphabet (A - Z) | 16.50 | 10.20 | 9.20 | 0.68 |
| Forms | 6.10 | 6.10 | 1.30 | 0.00 |
| Numbers (1 - 30) | 18.60 | 13.80 | 6.50 | 0.74 |
| Numbers Total | 32.10 | 28.60 | 6.30 | 0.56 |
| Relations | 14.30 | 13.30 | 2.30 | 0.43 |
| Sorting | 8.40 | 8.90 | 4.00 | -0.13 |
| Classification | 15.50 | 14.80 | 4.70 | 0.15 |
| Parts of Whole | 6.50 | 5.80 | 1.70 | 0.41 |
| Grand Total | 133.60 | 122.80 | 26.50 | 0.41 |
| PPVT Raw Score | 48.50 | 45.10 | 26.63 | 0.13 |
| PPVT IQ | 88.90 | 85.30 | 14.50 | 0.25 |
| Mean Age | 65.50 | 65.40 | 1.90 | 0.05 |

## 5. Australian Broadcasting Company Study (1972)

The report on this study was a preliminary one. Sesame Street was broadcast in Australia starting in January 1971. Unlike the Ball and Bogatz studies, the same children were not pretested and posttested. As an estimate of expected performance for their viewers, the Australian researchers used a group of nonviewers in the same locale the year before the series went on the air.

**Treatment and Design**

In December 1970, several Australian researchers selected a "fairly large but rather isolated centre of population" to conduct their study. The treatment was "exposure to Sesame Street" in its first season on the air. In December 1971, the researchers found that 84% of the children from 3 to 6 years of age had watched Sesame Street during the week in which they were interviewed. Only 5% of the children watched "irregularly" or not at all. They therefore validated the existence of the treatment in their study.

To implement their design, the researchers randomly selected half of the houses in the town for pretesting and the other half for posttesting. They asked the occupants to permit them to test all children between the ages of 3 and 6. At the pretest, 98% (N = 270) of the children participated. At the posttest, 99.75% (N = 394) participated. Thus, the design, in the Holland and Rubin model, used the performance of the previous year's children as an estimate of how the viewers would have performed had they not seen the Sesame Street programs.

**Analysis and Results**

The study included a number of outcome variables including the following:

1. Recognizing and naming some well known personalities from a children's television programme unrelated to Sesame Street.
2. Picture Vocabulary (IQ) Test
3. The Burt Graded Vocabulary Test
4. The Burt Oral Arithmetic Test
5. Discrimination Test (Numbers from Letters)
6. Identify (Name) printed letters and numbers when seen
7. Name printed shapes (square, circle, triangle, rectangle, star, diamond)
8. Counting Test

No data were given in the report, but the results were summarized in a narrative.

The two groups of children, viewers and nonviewers, did not differ on the Picture Vocabulary (IQ) Test - which the authors viewed as good evidence that the two groups were similar in ability as expected in random samples. The similarity held up for every age group.

The researchers reported "statistically" reliable differences in favor of the viewing group in the number of shapes correctly identified. Although positive in all age groups, the results were strongest in the "Age 4" group. In naming letters and numbers, the viewing group exceeded the nonviewing group at all age levels, but only the differences for the "Age 4" group were "statistically significant."

The researchers report that all other differences (Burt Graded Vocabulary Test, Burt Oral Arithmetic Test, and the Counting Test) were positive except in the case of "distinguishing printed letters from numbers." It is

not clear from the report what the nature of that task was. In any case, all differences but that one supported the assertion that Sesame Street had a positive impact on the children.

The final conclusion of the authors of the preliminary report was the following: "we have established with reasonable confidence that kids do learn some things from Sesame Street."

## 6. Ellis, Reid, and Hoen (July 1972)

Ellis, Reid, and Hoen studied the impact of Sesame Street on children in primary school in Vancouver.

**Treatment and Design**

In each of 8 randomly selected schools, one kindergarten class was randomly selected. In January 1972, the authors sent questionnaires to the children's parents asking them about the frequency with which their children viewed Sesame Street (0 through 5 times per week) in 1969-70, 1970-71, and 1971-72. The authors report that "95% of the kindergarten children watched Sesame Street" and "approximately two-thirds watched the program regularly." Using the parent responses, the researchers selected (method not specified) five "regular" or "high-frequency" (watched 3 to 5 times per week) viewers and five "irregular" or "low-frequency" (watched 0, 1, or 2 times per week) viewers in each class. Teachers gave performance tests to the two groups of children.

The treatment group (N=40) was the group of "regular" viewers of Sesame Street; the comparison group (N=40) was the group of "irregular" viewers.

Note that, in the Holland and Rubin sense, these authors are assuming that the performance of a group of 40 low-viewers provide a reasonable estimate of the performance of the 40 high-viewers in the absence of the program. The design makes sense to the degree that this assumption is reasonable.

**Analysis and Results**

The dependent variables, tasks, and/or measures are listed in the table of results. The basic performance measure was the percent of children who performed the task perfectly, i.e., answered all items correctly. The performance of the "high-viewers" was compared to that of the "low-viewers" on each of eleven outcome measures. The "high-viewers" outperformed the "low-viewers" on all eleven tasks. In 4 cases, the differences were statistically significant as shown in column three of the table.

"ELLIS, REID, and HOEN" - July 1972

**Study 1: High and Low Frequency Viewers in Vancouver**

| Variables | Treatment N=40 | Control N=40 | Effect |
|---|---|---|---|
| Capital Letters | 21 | 12 | +.005 |
| Small Letters | 14 | 8 | +.025 |
| Word with initial letters | 15 | 8 | + NS |
| Print Letters | 20 | 10 | + NS |
| Name Numerals | 27 | 19 | + .05 |
| Show Quantity of Numeral | 22 | 19 | + NS |
| Name Square | 38 | 34 | + NS |
| Name Cirle | 39 | 35 | + NS |
| Name Rectangle | 35 | 23 | + .01 |
| Name Triangle | 36 | 29 | + NS |
| Name round/rectangular objects | 33 | 25 | + NS |

## 7. Lukoff (June 1972)

This small study was conducted in two day care centers in New York City with four year old children.

**Treatment and Design**

In this case, one center scheduled Sesame Street programs three times a week. The second center did not. In the Holland/Rubin approach, T is the group (N = 14) in Center 1; C (N = 25) is the group in center 2.

**Analysis and Results**

Variables, measures, and/or tasks consisted of six brief tests - two matching, two recalling, and two recognizing varieties of letters, numbers, and geometric figures. Although the author did not make explicit that the comparison group was somehow an approximation for what the performance of the treatment group would have been had they not viewed Sesame Street several times a week, she attended to the issue by referring to previous research about the relative difficulties of the letters of the alphabet, by comparing the groups on letters "especially targeted" and "not targeted at all" by the Sesame Street shows. The inclusion of such information strengthened the validity of the argument and the credibility of this very small research study.

**Conclusion**

This study provides a simple, convincing argument for the positive impact of Sesame Street on the children's knowledge of the alphabet.

## 8. Minton (May 1972)

Judith Minton investigated the impact of Sesame Street's first season of programming on all kindergarten children in one school district in New York City.

**Treatment and Design**

In this research, the author used archival data available in the school district's central files. In the Holland/Rubin sense, the treatment group consisted of all the children in all the kindergarten classes in 1970. The series went on the air in November 1969 and completed one full series by the end of the 1969-70 school year. One comparison group included the children who completed kindergarten in 1969; a second comparison group included the children who completed kindergarten in 1968.

**Analysis and Results**

The school district administered the Metropolitan Readiness Test (MRT) to all children at the end of kindergarten. Variables measured by the MRT are listed in the table of results. Past research had demonstrated that student scores on standardized tests did not change much in schools and districts. The variable that showed the greatest change in this study was the knowledge of the alphabet. Sesame Street's impact on children's knowledge of the alphabet was the most consistent finding across all studies.

"MINTON" - April 1972

Study 1: Comparison of Spring 1970 Archival Data with 1969 Data

| Variables | Treatment N=524 | Control N=495 | SD | Effect Size |
|---|---|---|---|---|
| Word Meaning | 10.55 | 10.14 | 2.96 | 0.14 |
| Listening | 11.37 | 11.37 | 2.46 | 0.00 |
| Matching | 7.59 | 8.13 | 3.02 | -0.18 |
| Alphabet | 10.45 | 8.49 | 4.31 | 0.45 |
| Numbers | 14.43 | 14.48 | 4.98 | -0.01 |
| Copying | 7.80 | 7.69 | 4.01 | 0.03 |
| Total | 62.08 | 60.44 | 16.58 | 0.10 |

Study 2: Comparison of Spring 1970 Archival Data with 1968 Data

| Variables | Treatment N=524 | Control N=482 | SD | Effect Size |
|---|---|---|---|---|
| Word Meaning | 10.55 | 10.45 | 2.75 | 0.04 |
| Listening | 11.37 | 11.28 | 2.51 | 0.04 |
| Matching | 7.59 | 7.88 | 3.12 | -0.09 |
| Alphabet | 10.45 | 9.00 | 4.39 | 0.33 |
| Numbers | 14.43 | 14.36 | 5.02 | 0.01 |
| Copying | 7.80 | 8.52 | 3.85 | -0.19 |
| Total | 62.08 | 61.36 | 16.36 | 0.04 |

## 9. Salomon, Eglstein, Finkelstein, R., Finkelstein, I., Mintzberg, Malve, & Velner (1972)

Salomon and his associates investigated the impact of Sesame Street on cognitive performance, viewing behavior, and several additional variables in two groups of children in Israel - children in kindergarten and children in Grades 2 and 3. The same data are reanalyzed by Mintzberg (1973) in a later report.

**Treatment and Design**

Starting in the fall of 1971, Sesame Street was broadcast twice a week over a four month period. A narration in Hebrew was superimposed on the programs. Salomon conducted three independent studies - a school-based study of 317 children in kindergarten, Grade 2, and Grade 3; a clinical observation study of children's attention and involvement behavior; and, a controlled experiment with 75 second graders in which some children watched Sesame Street 8 hours a day, some watched 8 hours of adventure films, and some watched nothing. The school-based study involved educational variables relevant to this review. Four classes of children at each of the grade levels (Kindergarten, N=93; Grade 2, N=118; Grade 3, N=106) participated in the study. Two classes at each level were encouraged to watch Sesame Street at home, and two classes were not encouraged. As in the 1970 Ball and Bogatz study, children's viewing was not controlled, and the study did not include a nonviewing control group.

Children were asked about the amount of their viewing of "yesterday's" show on six occasions during the course of the study. From these data, an average viewing score was determined for each child. The children's reports were checked against parent reports of viewing. The correlation between children's reports and parent's reports was .72.

The parents of half the children were assembled twice and urged to watch the show with their children, to give appropriate explanations, and to encourage the children to watch the shows.

**Analysis and Results**

In these studies, Outcome or Dependent Variables, measures, and/or tasks included the following:

1. Population 1: Kindergarten
    (1) Letter Matching (5 items)
    (2) Number Matching (5 items)
    (3) Picture-Number Matching (12 items)
    (4) Relational Concepts (7 items)
    (5) Parts of the Whole (9 items)
    (6) Classification (18 items).
    (7) CEFT: An adaptation of Witkin's Children's Embedded Figure Test (Field Dependence/Independence)
    (8) Ordering of Pictures: children were asked to order in a logical way each of three sets of pictures.
2. Population 2: Grades 2 and 3
    (1) CEFT (as in #7 above)
    (2) Parts of Whole (#5 above)
    (3) Classification (#6 above)
    (4) Ordering of Pictures
    (5) Points of View
    (6) Figure and Ground
    (7) Close up; Long shot

At posttest, each child received a test which measured familiarity with the program: Sesame Street Retention Test (12 items).

The results showed that viewing and comprehension of the show's content were positively related and that the positive relationship increases during the first four months of the Salomon study ($r = .50$ at the beginning; $r = .76$ after four months).

To assess the impact of Sesame Street on posttest variables, Salomon and his associates regressed the posttest variables on SES, background variables (age, number of siblings, father's occupation), pretest variables, and amount of exposure to Sesame Street. At each grade level, exposure to Sesame Street resulted in a significant increase in the posttest variance accounted for by other independent variables in the study. Although the results support the conclusion that Sesame Street had a positive impact on the children, in the absence of well defined experimental and control conditions, the evidence is not conclusive.

## 10. Diaz-Guerrero, Reyes-Lagunes, Witzke, Holtzman (1973)

In this study, the authors investigated the impact of a completely new production of Sesame Street, Plaza Sesamo, on the learning skills of children in Mexico.

**Treatment and Design**

Preschool children (N = 221) in three day care centers provided the sample for this study. They were stratified by age and sex and randomly assigned to treatment and control groups. Children in the treatment group watched 130 Plaza Sesamo programs five days a week for six months. Children in the control group watched cartoons and other non-educational television programs. The design was a straightforward one, the control group children were the proxy for the expected performance of the treatment children in the absence of the treatment.

**Analysis and Results**

Diaz-Guerrero and his associates used adaptations of the ETS materials developed by Ball and Bogatz. The results are given in the tables on the following page. Assuming that the random assignment of children was valid, the comparisons of posttest scores and corresponding effect sizes are given in Study 1. The impact of Plaza Sesamo was positive and large across all variables.

Because the posttest measures are significantly dependent on the pretest measures, substantively and statistically, the authors also provide the results based on a regression of posttest scores on pretest scores, treatment, and a number of other variables of interest. In this analysis, the best estimate of the

treatment children's performance in the absence of the treatment is the performance of a hypothetical "control group" with the same initial scores as the treatment group. Using that comparison, the results in Study 2 show that Plaza Sesamo had a strong impact on all variables except "Parts of the Whole."

"DIAZ-GUERRERO" - October 1973
"DIAZ-GUERRERO and HOLTZMAN" - Spring 1974
"DIAZ-GUERRERO, REYES-LAGUNES, WITZKE, HOLTZMAN" - Spring 1976

Study 1: Children in 3 Day Care Centers in Mexico

| Variables | Treatment N=85 | Control N=88 | SD | Effect Size |
|---|---|---|---|---|
| Numbers | 27.90 | 21.50 | 7.10 | 0.90 |
| Letters/Words | 14.70 | 10.30 | 4.50 | 0.98 |
| Relations | 5.10 | 4.20 | 1.80 | 0.50 |
| Parts of the Whole | 5.20 | 4.20 | 2.30 | 0.43 |
| Ability fo Classify | 7.50 | 4.90 | 3.10 | 0.84 |
| Classification | 11.50 | 8.70 | 4.00 | 0.70 |
| Embedded Figures | 18.00 | 16.10 | 5.00 | 0.38 |
| Oral Comprehension | 21.60 | 17.90 | 6.10 | 0.61 |

Study 1: Same as Above - Impact in terms of Adjusted Measures

| Variables | Impact on 3yr olds | Impact on 4yr olds | Impact on 5yr olds |
|---|---|---|---|
| Numbers | 4.4 *** | 7.8 *** | 6.2 *** |
| Letters/Words | 2.9 *** | 4.5 *** | 5.1 *** |
| Relations | 0.5 | 1.6 ** | 0.7 |
| Parts of the Whole | 0.5 | 0.6 | 1.6 * |
| Ability to Sort | 0.5 | 3.0 *** | 3.2 *** |
| Classification | 2.0 ** | 3.2 *** | 2.3 * |
| Embedded Figures | 3.7 | 2.0 * | 0.5 |
| Oral Comprehension | 3.1 * | 5.3 ** | 2.4 * |
| General Knowledge | 2.7 | 7.3 *** | 4.8 *** |

Note: These results provide a "best" estimate for the net gain due to Sesame Street in the Holland/Rubin sense.

* $p <= .05$
** $p <= .01$
*** $p <= .001$

## 11. Lemercier and Teasdale (1973)

These researchers investigated the effectiveness of Sesame Street on a sample of children between the ages of 48 and 63 months in Australia. They did not reference the study conducted by the Australian Broadcasting Company discussed above.

**Treatment and Design**

The authors selected nine preschool kindergarten classes at random from a low SES area of a city. Children with disabilities or for whom English was not the basic language used at home were excluded from the study. The total number of children in the study was 67, 27 boys and 40 girls.

**Analysis and Results**

The authors' intent was to compare performance of high and low viewers on several tests from the ETS battery. They used familiarity with Sesame Street characters as an index of viewing intensity and divided the children into 32 "high viewers" and 35 "low viewers." In light of the Minton data given above, this assignment process is a weakness in the design. In the Minton data, large percentages of the three viewing groups - everyday, sometimes, never - identified 5 or 6 of the Sesame Street characters (44%, 43%, 30%). Those data suggest that 30% of the Australian "viewers" were not viewers at all. It is probably still true that the group identified as "high viewers" viewed somewhat more intensively than the group identified as "low viewers," but the variation in both groups is far from ideal for the Holland/Rubin approach.

"LEMERCIER and TEASDALE" - March 1973

Study 1: High and Low Frequency Viewers in Australia

| Variables | Treatment N=32 | Control N=35 | SD | Effect Size |
|---|---|---|---|---|
| Letters | 17.38 | 13.54 | 5.15 | 0.75 |
| Numbers | 29.00 | 21.23 | 7.72 | 1.01 |
| Sorting skills | 9.31 | 5.77 | 2.62 | 1.35 |
| Relations | 6.66 | 6.31 | 1.30 | 0.27 |
| Classification | 16.34 | 12.17 | 4.18 | 1.00 |
| Parts of the Whole | 6.47 | 5.23 | 2.44 | 0.51 |
| PPVT | 50.53 | 42.29 | 9.37 | 0.88 |

## 12. Darnell and Goodwin (1975)

In the spring of 1971, Darnell and Goodwin initiated a study of the kindergarten children (N = 450) in a sample of eight elementary schools in a small city in the Rocky Mountains. Sesame Street was not the main independent variable studied. However, the authors report that 98% of the children reported to have seen Sesame Street and 32% reported that they were consistent twice-a-day viewers during the previous two years (1969-70 and 1970-71 seasons).

**Treatment and Design**

Darnell and Goodwin examined the test data of 450 children who were in kindergarten in a small city near Denver in the 1970-71 school year (the "Kindergarten Child 1971" is the phrase they used). By the end of third grade, 122 children remained in the sample. The purpose of the researchers was to follow the children through their first three grades and investigate the influences of Head Start, other structured preschool experiences, and exposure to Sesame Street on their success in school. A treatment, in the Holland and Rubin sense, was not explicitly specified.

On entrance to kindergarten, the children were given the Metropolitan Readiness Test (MRT). At that time, the authors found that scores were correlated positively with family SES level, age, and sex. Amount of preschool experience was not correlated with score. Extent of Sesame Street viewing was correlated with some of the MRT scores - alphabet knowledge, listening, word meaning, numbers, and concepts. However, the authors pointed out that the relationship did not support the hypothesis that those who watched Sesame Street most functioned best. Rather, the relationship was curvilinear with moderate amounts of viewing appearing to relate most strongly with better performance on the MRT.

**Analysis and Results**

These authors reported that test results from 1968, 1972, and 1975 (same schools over time) indicated an overall increase of one-half to one standard deviation in average scores of the entering kindergarten classes. They saw those results as supporting a general conclusion of teachers and parents that the skill and cognitive levels of children at that age had been increasing over the period from 1968 through 1975. Unfortunately, they did not give the data in the report.

At the end of the first, second, and third grades, the children were given the Gates-MacGinitie Reading Test. Because the sample size decreased from 450 to 122, the researchers were concerned about differential attrition - dropout by children with special characteristics. They stated that overall mean scores and standard deviations on initial variables showed that the remaining sample was comparable to the original sample. That done, they regressed the children's reading scores on previous measures of reading, teacher judgments about the children's achievement, age, sex, family SES, and the extent of Sesame Street viewing and structured preschool experiences. They found "very little relationship" between "pre-kindergarten Sesame Street viewing" and reading success at the end of third grade.

In concluding their study, Darnell and Goodwin summarized their findings as follows:

> "The kindergarten child in this study was about five and one half years old, had attended preschool for one year, was a loyal viewer of Sesame Street, and came from a middle class family."

And, based on the results on the Metropolitan Readiness Tests, they drew the following final conclusion:

> "there was no consistent support for the hypothesis that children who attended preschool most (2 to 2 1/2 years) or who watched Sesame Street most (twice a day every day) functioned best on these school-related tasks."

## 13. Taylor and Skanes (1977)

Taylor and Skanes investigated the impact of Sesame Street on 5, 6, and 7 year old children living in isolated communities on the coast of Labrador. The authors hypothesized that because such children functioned 1 to 2 years below expected levels for children their age, they might profit from Sesame Street even though it was designed for younger children.

The population is unique - subsistence level fishing families. Most homes did not have central heating, sewage facilities, nor running water. The children had a restricted environment. They traveled little and had few books or educational materials in their homes. Education levels of parents were low and education was not highly valued.

**Treatment and Design**

Taylor and Skanes selected three samples of children from four communities. In community 1, broadcast was available - no control was exercised over viewing. In communities 2 and 3, no broadcast was available. In community 4, no broadcast was available - but Sesame Street was shown in kindergarten and first grade classes for 1 hour per day on a VCR. The programs were a part of the regular curriculum with the teacher interacting and facilitating interaction of the children with the program and with one another.

**Analysis and Results**

In this study, variables and measures were defined and collected by the following tests:

(1) Wechsler Preschool and Primary Scale of Intelligence (WPPSI),

(2) Peabody Picture Vocabulary Test (PPVT),

(3) Illinois Test of Psycholinguistic Abilities (ITPA), and

(4) Sesame Street Criterion Referenced Tests.

The authors examined repeated measures on these tests over a three year period. The results presented in the 1977 report were considered preliminary. Unfortunately, the authors only present the variables (test scores) which showed significant differences between the three community (viewing) groups. The results show that children who viewed Sesame Street gained significantly more than nonviewers on 7 of 24 CRTs related to the goals of the program. The authors state that by the end of Grade 2, all children scored at the top of the scale on the CRTs. They interpreted this as meaning that "children who see Sesame Street in school show initial but short lasting benefits."

"TAYLOR and SKANES" - 1977

**Study 1: Gain Scores in 3 Communities with varying access to Sesame Street**

| | Treatment 1<br>VCR Viewing | Treatment 2<br>TV available | No treatment<br>No TV or VCR |
|---|---|---|---|
| Letter sounds | 2.93 | 1.22 | 2.36 |
| Beginning sounds | 2.43 | 1.10 | 2.11 |
| Reading words | 2.10 | 1.48 | 2.07 |
| Word list | 2.60 | 1.76 | 2.59 |
| Sentences | 4.05 | 3.36 | 4.30 |
| Recitation between numbers | 4.55 | 4.36 | 5.57 |
| Label figures | 3.07 | 2.21 | 2.30 |
| Embedded figures | 3.43 | 2.85 | 3.02 |
| Inclusion | 3.67 | 3.45 | 2.73 |
| Auditory reception | 18.14 | 15.41 | 19.02 |
| Visual reception | 15.26 | 12.72 | 16.23 |
| Auditory sequential memory | 22.57 | 18.86 | 22.68 |
| Visual closure | 17.43 | 16.59 | 21.05 |
| Verbal expression | 14.69 | 11.09 | 13.21 |
| Grammatic closure | 12.71 | 9.93 | 12.07 |
| Sound blending | 16.07 | 11.16 | 16.25 |
| Mazes | 15.67 | 12.09 | 14.55 |
| PPVT | 39.48 | 40.76 | 44.84 |

## 14. Tower, Singer, Singer, and Biggs (1979)

Over a two week period, three groups of preschoolers were exposed to daily half-hour programs of Mister Rogers, Sesame Street, or a control series of nature and animal films. The authors investigated the effects of program format (Sesame Street, Mister Rogers, Nature/Animal films) on recall (only Sesame Street and Mister Rogers) and a variety of subsequent behaviors.

**Treatment and Design**

The authors randomly assigned children to one of the three conditions. They were pretested for IQ and "Imaginativeness." Initial data were examined to verify equivalent groups across the three conditions. Each day, children were questioned about material specific to the day's viewing. In the week following viewing, the spontaneous play behavior of the children was observed and assessed.

**Analysis and Results**

The recall study involved only Sesame Street and Mister Rogers. Ten questions per show were developed, 100 questions for Sesame Street, 100 questions for Mister Rogers. The researchers then selected five questions per show, attempting to balance questions across difficulty and shows. Questions were subdivided into those requiring memory of factual information and those requiring inference. Finally, the questions were sorted according to whether a concrete visual stimulus could be presented thus permitting visual recognition rather than recall. This approach resulted in the following: 28 factual and 22 inferential questions for Sesame Street; 27 factual and 23 inferential questions for Mister Rogers; 39 recall and 11 recognition questions for Sesame Street; and 41 recall and 9 recognition for Mister Rogers.

The results in terms of mean percentages of questions answered correctly are given on the following page. The results must be interpreted very tentatively in light of the fact that no pretesting of any of the questions were conducted. There is no way of knowing whether the questions in one treatment were more difficult in general than the questions in the other. The authors noted that the items used in the Mister Rogers part of the study had a greater range of difficulty. Nevertheless, the authors concluded, tentatively, that there was no overall difference (the totals were not significantly different). Children in the Mister Rogers group scored significantly higher on the factual memory and recognition items; the Sesame Street children scored significantly higher on the inferential items and higher (not statistically significant) on recall items.

"TOWER, SINGER, SINGER, and BIGGS" - April 1979

Study 1: Effects of Television Programming on Preschoolers' Cognition

| Variables | Sesame Street N=21 | SD | Mister Rogers N=21 | SD | Effect Size |
|---|---|---|---|---|---|
| Factual | 42.300 | 13.500 | 50.700 | 10.760 | -0.781 |
| Inferential | 48.100 | 16.570 | 39.000 | 10.330 | 0.881 |
| Recognition | 50.600 | 22.760 | 81.000 | 10.000 | -3.040 |
| Recall | 43.300 | 14.610 | 36.600 | 11.220 | 0.597 |
| Total | 44.700 | 13.770 | 45.500 | 9.250 | -0.086 |

## 15. Owens and Williams (1983)

Owens and Williams investigated the effects of viewing Sesame Street on learning styles and readiness skills on primary school children in a public school system in a rural area of Northeast Mississippi. They studied two groups of children. The first group of 15 preschool children watched an average of 5 hours of Sesame Street per week. The second group of 29 preschool children did not watch Sesame Street. The following review focuses on the readiness skills component of the study.

**Treatment and Design**

A sample of 44 primary school children (five and six year olds) in a rural area of northern Mississippi were identified. The Metropolitan Readiness Test (MRT) was administered to each child in August 1982. Parents were asked to indicate the number of hours per week the children spent watching Sesame Street. Those who watched five hours or more (N = 15) were placed Group 1. Those who did not watch Sesame Street at all (N = 29) were placed in Group 2. The treatment group was assumed to be viewing Sesame Street five hours or more per week. The group that did not watch Sesame Street was considered the comparison group; they represented what would have happened to Group 1 if they had not watched Sesame Street.

**Analysis and Results**

The authors computed the correlation of Sesame Street viewing (as a dichotomous variable - "no viewing" vs "more than five hours of viewing per week") with variables corresponding to the subtests of the Metropolitan Readiness Test. They found significant correlations between viewing and scores on the Word Meaning (r = .27), Alphabet (r = .21), and Matching (r = .21)

subtests. As in other studies, the authors did not attend to the problem of self-selection and the confounding of viewing and aptitude. Each could have been a cause of the other; each could have been caused by other related variables.

## 16. Field (1987)

Field investigated the effects on cognitive performance of coviewing television by five year old children and their parents. Viewing and coviewing data related to Sesame Street was not the central purpose of the study. However, portions of the data were used to estimate unique effects of viewing Sesame Street.

### Treatment and Design

Over a two and a half year period, 330 families, each with a five year old child, were studied. In two separate 10 day periods, parents were asked to keep diaries of all television viewing in the family from 6 A.M. until 2 A.M. the next morning. In general, the researchers found that children watched with others about 85% of the time. They investigated the effect of coviewing on cognitive performance by correlating viewing and coviewing measures with measures of cognitive performance.

### Analysis and Results

In general, the authors found that greater coviewing was negatively related to attention, comprehension, and reflectivity. On the other hand, amount of total viewing (whether coviewing or viewing alone) was positively related to "television comprehension," i.e., the recognition of television characters in common children's programs - including Sesame Street. Total viewing was not positively related to IQ measures used in the study. On the other hand, the authors report that "greater Sesame Street viewing additionally predicted higher IQ scores" ($r = .20$). Of course, as in other studies reviewed, viewing and aptitude are confounded because the children self-select their amount of viewing.